(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 November 2001 (08.11.2001)

PCT

(10) International Publication Number
WO 01/83749 A2

(51) International Patent Classification[7]: C12N 15/12, C07K 14/705, G01N 33/68, C12Q 1/68, C07K 16/28, A61K 38/17, A01K 67/027 // A61P 3/04, 3/10, 25/32

(21) International Application Number: PCT/US01/13387

(22) International Filing Date: 25 April 2001 (25.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/200,794	28 April 2000 (28.04.2000)	US
60/221,419	28 July 2000 (28.07.2000)	US
60/247,443	10 November 2000 (10.11.2000)	US

(71) Applicants *(for all designated States except US)*: **WARNER-LAMBERT COMPANY** [US/US]; 201 Tabor Road, Morris Plains, NJ 07950 (US). **THE MONELL CHEMICAL SENSES CENTER** [US/US]; 3500 Market Street, Philadelphia, PA 19104 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **BACHMANOV, Alexander, A.** [RU/US]; The Monell Chemical Senses Center, 3500 Market Street, Philadelphia, PA 19104 (US). **BEAUCHAMP, Gary, K.** [US/US]; The Monell Chemical Senses Center, 3500 Market Street, Philadelphia, PA 19104 (US). **CHATTERJEE, Aurobindo** [IN/US]; 2106 Clinton Avenue, Apt. A, Alameda, CA 94501 (US). **DE JONG, Pieter, J.** [US/US]; 631 Catamaran Street #4, Foster City, CA 94404 (US). **LI, Shanru** [CN/US]; The Monell Chemical Senses Center, 3500 Market Street, Philadelphia, PA 19104 (US). **Li, Xia** [CN/US]; The Monell Chemical Senses Center, 3500 Market Street, Philadelphia, PA 19104 (US). **OHMEN, Jeffrey, D.** [US/US]; 1817 Versailles Avenue, Alameda, CA 94501 (US). **REED, Danielle, R.** [US/US]; The Monell Chemical Senses Center, 3500 Market Street, Philadelphia, PA 19104 (US). **ROSS, David** [US/US]; 2651 Barbers Point Road, Apt. D., Alameda, CA 94501 (US). **TORDOFF, Michael, Guy** [GB/US]; The Monell Chemical Senses Center, 3500 Market Street, Philadelphia, PA 19104 (US).

(74) Agents: **FEDERMAN, Evan, J.**; Warner-Lambert Company, 201 Tabor Road, Morris Plains, NJ 07950 et al. (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/83749 A2

(54) Title: GENE AND SEQUENCE VARIATION ASSOCIATED WITH SENSING CARBOHYDRATE COMPOUNDS AND OTHER SWEETENERS

(57) Abstract: The present invention relates to the discovery of a gene and its sequence variation associated with preference for carbohydrates, other sweet compounds, or ethanol. The present invention also relates to the study of metabolic pathways to identify other genes, receptors, and relationships that contribute to differences in sensing of carbohydrates or ethanol. The present invention also relates to germline or somatic sequence variations and its use in the diagnosis and prognosis of predisposition to diabetes, other obesity related disorders, or ethanol consumption. The present invention also provided probes or primers specific for the detection and analysis of such sequence variation. The present invention also relates to method for screening drugs for inhibition or restoration of gene function as antidiabetic, antiobesity, or antialcohol consumption therapies. The present invention relates to other antidiabetic, antiobesity disorder, or antialcohol consumption therapies, such as gene therapy, protein replacement therapy, etc. Finally, the present invention relates to a method for identifying sweeteners or alcohols utilizing the gene and its variations.

GENE AND SEQUENCE VARIATION ASSOCIATED WITH SENSING CARBOHYDRATE COMPOUNDS AND OTHER SWEETENERS

FIELD OF THE INVENTION

The present invention relates generally to the field of mouse and human genetics and sensing of extracellular carbohydrates. Specifically, the present invention relates to the discovery of a gene and its sequence variation associated with a differential preference for sweet compounds in laboratory strains of mice.

BACKGROUND OF THE INVENTION

The ability to sense extra-cellular carbohydrates, transduce this sensory information, and relay it to the brain, is carried out by membrane bound receptors in taste papillae. Many approaches to identify the sweet receptor or receptors have been tried, but the problem has proved, until recently, to be difficult.

Mammals vary in their *ad libitum* consumption of sweeteners. To investigate the genetic contribution to this complex behavior, behavioral, electrophysiological, and genetic studies were conducted using two strains of mice that differ markedly in their preference for sucrose and saccharin (Bachmanov et al., *Behavior Genetics,* 1996;26:563-573).

Recently published data indicates that the ability to sense carbohydrates is linked to obesity. These studies demonstrated that sensation of simple carbohydrates is suppressible by the adipose hormone, leptin.

These studies demonstrated that a locus on the telomere of mouse chromosome 4 accounts for ~40% of the genetic variability in sucrose and saccharin intake, and that the effect of this locus is to enhance or retard the gustatory neural response to sucrose.

SUMMARY OF THE INVENTION

The present invention provides a gene and its sequence variation associated with a preference for carbohydrate compounds, other sweeteners, or alcohol.

The present invention provides a gene and its sequence variation associated a differential response by the pancreas and/or muscle in response to dietary carbohydrates.

The present invention also relates to sequence variation and its use in the diagnosis and prognosis of predisposition to diabetes, other obesity-related disorders, or alcohol consumption.

The present invention also relates to the study of taste to identify molecules responsible for signal transduction, other receptors and genes and relationships that contribute to taste preference.

The present invention also relates to the study of diabetes to identify molecules responsible for sensing extra-cellular carbohydrate, other receptors and genes and relationships that contribute to a diabetic state.

The present invention also relates to a sequence variation and its use in the identification of specific alleles altered in their specificity for carbohydrate compounds.

The present invention also relates to a recombinant construct comprising SAC1 (also referred to as *Sac*) polynucleotide suitable for expression in a transformed host cell.

The present invention also provides primers and probes specific for the detection and analysis of the SAC1 locus.

The present invention also relates to kits for detecting a polynucleotide comprising a portion of the SAC1 locus.

The present invention also relates to transgenic animals, which carry an altered SAC1 allele, such as a knockout mouse.

The present invention also relates to methods for screening drugs for inhibition or restoration of SAC1 function as a taste receptor.

The present invention also relates to identification of sweeteners or alcohols using the SAC1 gene and its sequence variations.

The present invention also relates to methods for screening drugs for inhibition or restoration of SAC1 function in homeostatic regulation of glucose levels.

The present invention also relates to methods for screening drugs for modification of SAC1 function in the consumption of alcohol.

Finally, the present invention provides therapies directed to diabetic or obesity disorders. Therapies of diabetes and obesity include gene therapy, protein replacement, protein mimetics, and inhibitors.

BRIEF DESCRIPTION OF THE FIGURES

Fig. 1A shows genetic mapping of the SAC1 locus, using 632 F2 mice from a cross between the B6 (high preference) and 129 (low preference) strains. Mapping results were obtained with MAPMAKER/QTL Version 1.1, using an unconstrained model. A black triangle at the bottom indicates peak LOD score at M134G01 marker. Horizontal line at the bottom shows a 1-LOD confidence interval.

Fig. 1B shows SAC1-containing chromosomal region defined by a donor fragment of the 129.B6-*Sac*^{b} partially congenic mice. The partially congenic strains were constructed by identifying several founder F2 mice with small fragments of the telomeric region of mouse chromosome 4 from the B6 strain and successive backcrossing to the 129 strain. Presence and size of donor fragment were determined by genotyping polymorphic markers in mice from the N4, N6, N7, N4F4, and N3F5 generations.

Fig. 1C shows average daily saccharin consumption by N6, N7, N4F4, and N3F5 segregating partially congenic 129.B6-*Sac* mice in 4-days two-bottle tests with water (means ± SE). The open bar indicates intakes of mice that did not inherit the donor fragment. The black bar indicates intakes of mice with one or two copies of the donor fragment, which is flanked by *280G12-T7* proximally and

D4Mon1 distally. The complete donor fragment is represented by overlapping sequences of the BAC RPCI-23-118E21 and a genomic clone (Accession AF185591), as indicated at the bottom. The size of the SAC1-containing donor fragment is 194, 478 kb.

Fig. 1D shows BAC contig of distal chromosome 4 in the SAC1 region. Using ^{32}P radioactively labeled probes from the nonrecombinant interval, a mouse BAC library (RPCI-23) was screened; positive clones were confirmed by PCR analysis and only clones positive by hybridization and by PCR are included in the contig. BAC ends were sequenced and PCR primers designed. The STS content of each BAC, using all BAC ends was determined. BAC size was determined by digesting the BAC with *NotI*, and the insert size determined using pulse field gel electrophoresis.

Fig. 1E shows genes contained within the SAC1 nonrecombinant interval. Arrows indicate predicted direction of transcription. See Table 1 for a description of gene prediction, and details concerning function.

Fig. 2A shows the mouse SAC1 gene (mSac; Accession AF311386), its human ortholog (hSac), and the previously described gene T1R1, now Gpr70, are aligned above. Residues shaded in black are identical between at least two identical residues; residues in gray indicate conservative changes. The human ortholog was identified by sequence homology search within the *htgs* database (Accession AC026283). The amino acid sequence of the human ortholog was predicted using GENSCAN. The amino acid sequence of mouse *Gpr70* was obtained by constructing primers based upon the nucleotide sequence, and taste cDNA was amplified and sequenced. This amino acid and nucleotide sequence for *Gpr70* differed slightly from the initial report; the sequence reported in this paper has been deposited in GenBank (AF301161, AF301162). The location of the missense mutation is indicated by an *.

Fig. 2B shows structure of the SAC1 gene. The six exons are shown as black boxes.

Fig. 2C shows conformation of a protein predicted from the *Sac* gene. To determine the transmembrane regions, the hydrophobicity was determined using

the computer program HMMTOP, and drawn with TOPO. The missense mutation is denoted with an asterisk.

Fig. 3 shows saccharin and sucrose preferences by mice from inbred strains with two different haplotypes of the *Sac* gene. The haplotype found in the B6 mice and the other high sweetener-preferring inbred strains consisted of four variants, two variants were 5′ of the predicted translation start codon, one variant was a missense mutation (Ile61Thr), and the last variant was located in the intron between exon 2 and 3. The strains with the B6-like haplotype of *Sac* strongly preferred saccharin (82 ± 4%) and sucrose (86 ± 6%), whereas strains with the 129-like haplotype were indifferent to these solutions (57 ± 2% and 54 ± 1% respectively, $p = 0.0015$).

Fig. 4A shows tissue expression of the SAC1 gene. Note that cDNA was obtained from a commercial source for the multiple tissue panel, with the exception of tongue cDNA, which was as isolated by the investigator, as described within the text. Relative band intensities may differ due to differences in cDNA isolation methods or concentration.

Fig. 4B shows RNA from human fungiform papillae was obtained from biopsy material, reversed transcribed, and the resulting bands from genomic and cDNA were amplified using primers, described in the text. The bands were excised from the agarose gel, purified and reamplified. The PCR product was sequenced to confirm that the bands amplified the human otholog to *Sac*.

Fig. 5 shows amino acid sequence alignment of the mouse cDNA sequence for the SAC1 gene and the cDNA for a calcium sensing metabotropic receptor. Dark areas indicated regions of shared similarity.

Fig. 6 plots the hydrophobicity of the SAC1 amino acid sequence as predicted by the computer program Top Pred. Note the seven transmembrane domains characteristic of G-protein coupled receptors.

DETAILED DESCRIPTION OF THE INVENTION

I. Definitions

The present invention employs the following definitions:

As used herein, the terms "polynucleotide" and "nucleic acid" refer to naturally occurring polynucleotides, e.g., DNA or RNA. These terms do not refer to a specific length. Thus, these terms include oligonucleotide, primer, probe, etc. These terms also refer to analogs of naturally occurring polynucleotides. The polynucleotide may be double stranded or single stranded. The polynucleotides may be labeled with radiolabels, fluorescent labels, enzymatic labels, proteins, haptens, antibodies, sequence tags.

For example, these terms include RNA, cDNA, genomic DNA, synthetic forms, and mixed polymers, both sense and antisense strands, and may be chemically or biochemically modified or may contain non-natural or derivatized nucleotide bases, as will be readily appreciated by those skilled in the art. Such modifications include, for example, labels, methylation, substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoramidates, carbamates, etc.), charged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), pendent moieties (e.g., polypeptides), intercalators (e.g., acridine, psoralen, etc.), chelators, alkylators, and modified linkages (e.g., alpha anomeric nucleic acids, etc.). Also included are synthetic molecules that mimic polynucleotides in their ability to bind to a designated sequence via hydrogen bonding and other chemical interactions. Such molecules are known in the art and include, for example, those in which peptide linkages substitute for phosphate linkages in the backbone of the molecule.

As used herein, the term "polynucleotide amplification" refers to a broad range of techniques for increasing the number of copies of specific polynucleotide sequences. Typically, amplification of either or both strand of the target nucleic acid comprises the use of one or more nucleic acid-modifying enzymes, such as a DNA polymerase, a ligase, an RNA polymerase, or an RNA-dependent reverse transcriptase. Examples of polynucleotide amplification reaction include, but not

limited to, polymerase chain reaction (PCR), nucleic acid sequence based amplification (NASB), self-sustained sequence replication (3SR), strand displacement activation (SDA), ligase chain reaction (LCR), Qβ replicase system, and the like.

As used herein, the term "primer" refers to a nucleic acid, e.g., synthetic polynucleotide, which is capable of annealing to a complementary template nucleic acid (e.g., the SAC1 locus) and serving as a point of initiation for template-directed nucleic acid synthesis. A primer need not reflect the exact sequence of the template but must be sufficiently complementary to hybridize with a template. Typically, a primer will include a free hydroxyl group at the 3′ end. The appropriate length of a primer depends on the intended use of the primer but typically ranges from 12 to 30 nucleotides. The term primer pair means a set of primers including a 5′ upstream primer that hybridizes with the 5′ end of the target sequence to be amplified and a 3′ downstream primer that hybridizes with the complement of the 3′ end of the target sequence to be amplified.

The present invention includes all novel primers having at least eight nucleotides derived from the SAC1 locus for amplifying the SAC1 gene, its complement or functionally equivalent nucleic acid sequences. The present invention does not include primers which exist in the prior art. That is, the present invention includes all primers having at least 8 nucleotides with the proviso that it does not include primers existing in the prior art.

"Target polynucleotide" refers to a single- or double-stranded polynucleotide which is suspected of containing a target sequence, and which may be present in a variety of types of samples, including biological samples.

"Antibody" refers to polyclonal and/or monoclonal antibody and fragments thereof, and immunologic binding equivalents thereof, which are capable of specifically binding to the SAC1 polypeptides and fragments thereof or to polynucleotide sequences from the SAC1 region, particularly from the SAC1 locus or a portion thereof. Antibody may be a homogeneous molecular entity, or a mixture such as a serum product made up of a plurality of different molecular entities.

Antibodies may be produced by in vitro or in vivo techniques well-known in the art. For example, for production of polyclonal antibodies, an appropriate target immune system, typically mouse or rabbit, is selected. Substantially purified antigen is presented to the immune system. Typical sites for injection are in footpads, intramuscularly, intraperitoneally, or intradermally. Polyclonal antibodies may then be purified and tested for immunological response, e.g., using an immunoassay.

For production of monoclonal antibodies, protein, polypeptide, fusion protein, or fragments thereof may be injected into mice. After the appropriate period of time, the spleens may be excised and individual spleen cells fused, typically, to immortalized myeloma cells under appropriate selection conditions. Thereafter, the cells are clonally separated and the supernatants of each clone tested for their production of an appropriate antibody specific for the desired region of the antigen. Affinities of monoclonal antibodies are typically 10^{-8} M^{-1} or preferably 10^{-9} to 10^{-10} M^{-1} or stronger.

Other suitable techniques involve in vitro exposure of lymphocytes to the antigenic polypeptides, or alternatively, to selection of libraries of antibodies in phage or similar vectors.

Frequently, antibodies are labeled by joining, either covalently or non-covalently, a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent agents, chemiluminescent agents, magnetic particles, and the like. Also, recombinant immunoglobulins may be produced.

"Binding partner" refers to a molecule capable of binding another molecule with specificity, as for example, an antigen and an antigen-specific antibody or an enzyme and its inhibitor. Binding partners are known in the art and include, for example, biotin and avidin or streptavidin, IgG and protein A, receptor-ligand couples, and complementary polynucleotide strands. In the case of complementary polynucleotide binding partners, the partners are normally at least about 15, 20, 25, 30, 40 bases in length.

A "biological sample" refers to a sample of tissue or fluid suspected of containing an analyte (e.g., polynucleotide, polypeptide) including, but not limited to, e.g., plasma, serum, spinal fluid, lymph fluid, the external sections of the skin, respiratory, intestinal, and genitourinary tracts, tears, saliva, blood cells, organs, tissue and samples of in vitro cell culture constituents. A biological sample is typically from human or other animal.

"Encode." A polynucleotide is said to "encode" a polypeptide if, in its native state or when manipulated by methods well-known to those skilled in the art, it can be transcribed and/or translated to produce the mRNA and/or the polypeptide or a fragment thereof. The antisense strand is the complement of such a nucleic acid, and the encoding sequence can be deduced therefrom.

"Isolated" or "substantially pure" polynucleotide or polypeptide (e.g., an RNA, DNA, protein) is one which is substantially separated from other cellular components which naturally accompany a native human nucleic acid or protein, e.g., ribosomes, polymerases, many other human genome sequences and proteins. The term embraces a nucleic acid or peptide sequence which has been removed from its naturally occurring environment, and includes recombinant or cloned DNA isolates and chemically synthesized analogs or analogs biologically synthesized by heterologous systems.

"SAC1 Allele" refers to normal alleles of the SAC1 locus as well as alleles carrying variations that predispose individuals to develop obesity, diabetes, or for alcohol consumption or alcoholism.

"SAC1 Locus" refers to polynucleotides, which are in the SAC1 region, that are likely to be expressed in normal individual, certain alleles of which predispose an individual to develop obesity, diabetes, or alcohol consumption or alcoholism. The SAC1 locus includes coding sequences, intervening sequences and regulatory elements controlling transcription and/or translation. The SAC1 locus includes all allelic variations of the DNA sequence.

The DNA sequences used in this invention will usually comprise at least about 5 codons (15 nucleotides), 7, 10, 15, 20, or 30 codons, and most preferably, at least about 35 codons. One or more introns may also be present. This number of

nucleotides is usually about the minimal length required for a successful probe that would hybridize specifically with a SAC1 locus.

"SAC1 Region" refers to a portion of mouse chromosome 4 bounded by the markers 280G12-T7 and D4Mon1 GenBank Accession number is YG7772 (SEQ ID NO: 652) and is GCAGTGAGCTGCAGAGTTTGCAGAATGAGGGCACTCTAAACTCATCAA GTGAGGAGGCCCTTCCCTCACACTCCAGATGGCTGATAGGTGGCATTA CATGGTC(CA)nCGCGCGCACGCGCTCAGATGCAATCTCCACATTCATA ACCAGATGTCCTTGGGTAGGCCT. The CA sequence in the middle is variable in length. In the B6 mouse, n = 19, while in the 129 mouse, n = 16. This region contains the SAC1 locus, including the SAC1 gene. GenBank accession number for the SAC1 gene is AF311386.

As used herein, a "portion" or "fragment" of the SAC1 gene, locus, region, or allele is defined as having a minimal size of at least about 15 nucleotides, or preferably at least about 20, or more preferably at least about 25 nucleotides, and may have a minimal size of at least about 40 nucleotides.

As used herein, the term "polypeptide" refers to a polymer of amino acids without referring to a specific length. This term includes to naturally occurring protein. The term also refers to modifications, analogues and functional mimetics thereof. For example, modifications of the polypeptide may include glycosylations, acetylations, phosphorylations, and the like. Analogues of polypeptide include unnatural amino acid, substituted linkage, etc. Also included are polypeptides encoded by DNA which hybridize under high or low stringency conditions, to the nucleic acids of interest.

Modification of polypeptides includes those substantially homologous to primary structural sequence, e.g., in vivo or in vitro chemical and biochemical modifications or incorporation unusual amino acids. Such modifications include, for example, acetylation, carboxylation, phosphorylation, glycosylation, ubiquitination, labeling, e.g., with radionuclides, and various enzymatic modifications, as will be readily appreciated by those well-skilled in the art. A variety of methods for labeling polypeptides and of substituents or labels useful for such purposes are well-known in the art, and include radioactive isotopes such

as ^{32}P, ligands which bind to labeled antiligands (e.g., antibodies), fluorophores, chemiluminescent agents, enzymes, and antiligands which can serve as specific binding pair members for a labeled ligand. The choice of label depends on the sensitivity required, ease of conjugation with the primer, stability requirements, and available instrumentation. Methods of labeling polypeptides are well-known in the art (see Sambrook et al., 1989 or Ausubel et al., 1992).

Besides substantially full-length polypeptides, the present invention provides for biologically active fragments of the polypeptides. Significant biological activities include ligand-binding, immunological activity, and other biological activities characteristic of SAC1 polypeptides. Immunological activities include both immunogenic function in a target immune system, as well as sharing of immunological epitopes for binding, serving as either a competitor or substitute antigen for an epitope of the SAC1 protein. As used herein, "epitope" refers to an antigenic determinant of a polypeptide. An epitope could comprise three amino acids in a spatial conformation that is unique to the epitope. Generally, an epitope consists of at least five such amino acids, and more usually consists of at least 8 to 10 such amino acids. Methods of determining the spatial conformation of such amino acids are known in the art.

For immunological purposes, tandem-repeat polypeptide segments may be used as immunogens, thereby producing highly antigenic proteins. Alternatively, such polypeptides will serve as highly efficient competitors for specific binding.

Fusion proteins comprise SAC1 polypeptides and fragments. Homologous polypeptides may be fusions between two or more SAC1 polypeptide sequences or between the sequences of SAC1 and a related protein. Likewise, heterologous fusions may be constructed which would exhibit a combination of properties or activities of the derivative proteins. For example, ligand-binding or other domains may be "swapped" between different new fusion polypeptides or fragments. Such homologous or heterologous fusion polypeptides may display, for example, altered strength or specificity of binding. Fusion partners include immunoglobulins, bacterial β-galactosidase, trpE, protein A, β-lactamase, α-amylase, alcohol dehydrogenase, and yeast α mating factor.

Fusion proteins will typically be made by either recombinant nucleic acid methods or may be chemically synthesized. Techniques for the synthesis of polypeptides are known in the art.

Functional mimetics of a native polypeptide may be obtained using known methods in the art. For example, polypeptides may be least about 50% homologous to the native amino acid sequence, preferably in excess of about 70%, and more preferably at least about 90% homologous. Substitutions typically contain the exchange of one amino acid for another at one or more sites within the polypeptide, and may be designed to modulate one or more properties of the polypeptide, such as stability against proteolytic cleavage, without the loss of other functions or properties. Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. Preferred substitutions are ones which are conservative, that is, one amino acid is replaced with one of similar shape and charge. Conservative substitutions are well-known in the art and typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and tyrosine, phenylalanine.

Certain amino acids may be substituted for other amino acids in a polypeptide structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules or binding sites on proteins interacting with a polypeptide. Since it is the interactive capacity and nature of a polypeptide which defines that polypeptide's biological functional activity, certain amino acid substitutions can be made in a protein sequence, and its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydrophobic amino acid index in conferring interactive biological function on a protein is generally understood in the art. Alternatively, the substitution of like amino acids can be made effectively on the basis of hydrophilicity.

A peptide mimetic may be a peptide-containing molecule that mimics elements of protein secondary structure. The underlying rationale behind the use of peptide mimetics is that the peptide backbone of proteins exists chiefly to orient amino acid side chains in such a way as to facilitate molecular interactions, such as those of antibody and antigen, enzyme and substrate or scaffolding proteins. A peptide mimetic is designed to permit molecular interactions similar to the natural molecule. A mimetic may not be a peptide at all, but it will retain the essential biological activity of a natural polypeptide.

Polypeptides may be produced by expression in a prokaryotic cell or produced synthetically. These polypeptides typically lack native post-translational processing, such as glycosylation. Polypeptides may be labeled with radiolabels, fluorescent labels, enzymatic labels, proteins, haptens, antibodies, sequence tags.

"SAC1 polypeptide" refers to a protein or polypeptide encoded by the SAC1 locus, variants, fragments or functional mimics thereof. A SAC polypeptide may be that derived from any of the exons described herein which may be in isolated and/or purified form. The length of SAC1 polypeptide sequences is generally at least about 5 amino acids, usually at least about 10, 15, 20, 30 residues.

"Alcohol consumption" relates to the intake and/or preference of an animal for ethanol.

"Diabetes" refers to any disorder that exhibits phenotypic features of an increased or decreased level of a biological substance associated with glucose or fatty acid metabolism. The term "carbohydrate" refers to simple mono and disaccharides.

The terms "sequence variation" or "variant form" encompass all forms of polymorphism and mutations. A sequence variation may range from a single nucleotide variation to the insertion, modification, or deletion of more than one nucleotide. A sequence variation may be located at the exon, intron, or regulatory region of a gene.

Polymorphism refers to the occurrence of two or more genetically determined alternative sequences or alleles in a population. A biallelic polymorphism has two forms. A triallelic polymorphism has three forms. A

polymorphic site is the locus at which sequence divergence occurs. Diploid organisms may be homozygous or heterozygous for allelic forms. Polymorphic sites have at least two alleles, each occurring at frequency of greater than 1% of a selected population. Polymorphic sites also include restriction fragment length polymorphisms, variable number of tandem repeats (VNTRs), hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats, simple sequence repeats, and insertion elements. The first identified allelic form may be arbitrarily designated as the reference sequence and other allelic forms may be designated as alternative or variant alleles. The allelic form occurring most frequently in a selected population is sometimes referred to as the wild type form or the consensus sequence.

Mutations include deletions, insertions and point mutations in the coding and noncoding regions. Deletions may be of the entire gene or of only a portion of the gene. Point mutations may result in stop codons, frameshift mutations, or amino acid substitutions. Somatic mutations are those which occur only in certain tissues, such as liver, heart, etc. and are not inherited in the germline. Germline mutations can be found in any of a body's tissues and are inherited.

"Operably linked" refers to a juxtaposition wherein the components are in a relationship permitting them to function in their intended manner. For instance, a promoter is operably linked to a coding sequence if the promoter affects its transcription or expression.

The term "probes" refers to polynucleotide of any suitable length which allows specific hybridization to the target region. Probes may be attached to a label or reporter molecule using known methods in the art. Probes may be selected by using homologous polynucleotides. Alternatively, polynucleotides encoding these or similar polypeptides may be synthesized or selected by use of the redundancy in the genetic code. Various codon substitutions may be introduced, e.g., by silent changes (thereby producing various restriction sites) or to optimize expression for a particular system. Mutations may be introduced to modify the properties of the polypeptide, perhaps to change ligand-binding affinities, interchain affinities, or the polypeptide degradation or turnover rate.

Probes comprising synthetic oligonucleotides or other polynucleotides of the present invention may be derived from naturally occurring or recombinant single- or double-stranded polynucleotides, or be chemically synthesized. Probes may also be labeled by nick translation, Klenow fill-in reaction, or other methods known in the art.

Portions of the polynucleotide sequence having at least about 8 nucleotides, usually at least about 15 nucleotides, and fewer than about 6 kb, usually fewer than about 1.0 kb, from a polynucleotide sequence encoding SAC1 are preferred as probes.

The terms "isolated," "substantially pure," and "substantially homogeneous" are used interchangeably to describe a protein or polypeptide which has been separated from components which accompany it in its natural state. A monomeric protein is substantially pure when at least about 60% to 75% of a sample exhibits a single polypeptide sequence. A substantially pure protein will typically comprise about 60% to 90% W/W of a protein sample, more usually about 95%, and preferably will be over about 99% pure. Protein purity or homogeneity may be indicated by a number of means well-known in the art, such as polyacrylamide gel electrophoresis of a protein sample, followed by visualizing a single polypeptide band upon staining the gel. For certain purposes, higher resolution may be provided by using HPLC or other means well-known in the art which are utilized for purification.

A SAC1 protein is substantially free of naturally associated components when it is separated from the native contaminants which accompany it in its natural state. Thus, a polypeptide which is chemically synthesized or synthesized in a cellular system different from the cell from which it naturally originates will be substantially free from its naturally associated components. A protein may also be rendered substantially free of naturally associated components by isolation, using protein purification techniques well-known in the art.

"Recombinant nucleic acid" is a nucleic acid which is not naturally occurring, or which is made by the artificial combination of two otherwise separated segments of sequence. This artificial combination is often accomplished by either chemical synthesis means, or by the artificial manipulation of isolated

segments of nucleic acids, e.g., by genetic engineering techniques. Such is usually done to replace a codon with a redundant codon encoding the same or a conservative amino acid, while typically introducing or removing a sequence recognition site. Alternatively, it is performed to join together nucleic acid segments of desired functions to generate a desired combination of functions.

"Regulatory sequences" refers to those sequences normally within 100 kb of the coding region of a locus, but they may also be more distant from the coding region, which affect the expression of the gene (including transcription of the gene, and translation, splicing, stability or the like of the messenger RNA).

"Substantial homology or similarity." A nucleic acid or fragment thereof is of substantially homologous ("or substantially similar") to another if, when optimally aligned (with appropriate nucleotide insertions or deletions) with the other nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least about 60% of the nucleotide bases, usually at least about 70%, more usually at least about 80%, preferably at least about 90%, and more preferably at least about 95-98% of the nucleotide bases.

Identity means the degree of sequence relatedness between two polypeptide or two polynucleotides sequences as determined by the identity of the match between two strings of such sequences. Identity can be readily calculated (Lesk A.M., ed., *Computational Molecular Biology*, New York: Oxford University Press, 1988; Smith D.W., ed., *Biocomputing: Informatics and Genome Projects*, New York: Academic Press, New York, 1993; Griffin A.M., and Griffin H.G., eds., *Computer Analysis of Sequence Data*, Part 1, New Jersey: Humana Press, 1994; von Heinje G., *Sequence Analysis in Molecular Biology*, Academic Press, 1987; and Gribskov M. and Devereux J., eds., *Sequence Analysis Primer*, New York: M Stockton Press, 1991).

Alternatively, substantial homology or similarity exists when a nucleic acid or fragment thereof will hybridize to another nucleic acid (or a complementary strand thereof) under selective hybridization conditions, to a strand, or to its complement. Selectivity of hybridization exists when hybridization which is substantially more selective than total lack of specificity occurs. Typically, selective hybridization will occur when there is at least about

55% homology over a stretch of at least about 14 nucleotides, preferably at least about 65%, more preferably at least about 75%, and most preferably at least about 90%. The length of homology comparison, as described, may be over longer stretches, and in certain embodiments will often be over a stretch of at least about 9 nucleotides, usually at least about 20 nucleotides, more usually at least about 24 nucleotides, typically at least about 28 nucleotides, more typically at least about 32 nucleotides, and preferably at least about 36 or more nucleotides.

Nucleic acid hybridization will be affected by such conditions as salt concentration, temperature, or organic solvents, in addition to the base composition, length of the complementary strands, and the number of nucleotide base mismatches between the hybridizing nucleic acids, as will be readily appreciated by those skilled in the art. Stringent temperature conditions will generally include temperatures in excess of 30°C, typically in excess of 37°C, and preferably in excess of 45°C. Stringent salt conditions will ordinarily be less than 1000 mM, typically less than 500 mM, and preferably less than 200 mM. However, the combination of parameters is much more important than the measure of any single parameter.

The terms "substantial homology" or "substantial identity," when referring to polypeptides, indicate that the polypeptide or protein in question exhibits at least about 30% identity with an entire naturally-occurring protein or a portion thereof, usually at least about 70% identity, and preferably at least about 95% identity.

Homology, for polypeptides, is typically measured using sequence analysis software (see, e.g., the Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center). Protein analysis software matches similar sequences using measures of homology assigned to various substitutions, deletions and other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine.

"Substantially similar function" refers to the function of a modified nucleic acid or a modified protein, with reference to the wild-type SAC1 nucleic acid or

wild-type SAC1 polypeptide. The modified polypeptide will be substantially homologous to the wild-type SAC1 polypeptide and will have substantially the same function. The modified polypeptide may have an altered amino acid sequence and/or may contain modified amino acids. In addition to the similarity of function, the modified polypeptide may have other useful properties, such as a longer half-life. The similarity of function (activity) of the modified polypeptide may be substantially the same as the activity of the wild-type SAC1 polypeptide. Alternatively, the similarity of function (activity) of the modified polypeptide may be higher than the activity of the wild-type SAC1 polypeptide. The modified polypeptide is synthesized using conventional techniques, or is encoded by a modified nucleic acid and produced using conventional techniques. The modified nucleic acid is prepared by conventional techniques. A nucleic acid with a function substantially similar to the wild-type SAC1 gene function produces the modified protein described above.

A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 to 7 contiguous amino acids, often at least about 7 to 9 contiguous amino acids, typically at least about 9 to 13 contiguous amino acids and, most preferably, at least about 20 to 30 or more contiguous amino acids.

The polypeptides of the present invention, if soluble, may be coupled to a solid-phase support, e.g., nitrocellulose, nylon, column packing materials (e.g., Sepharose beads), magnetic beads, glass wool, plastic, metal, polymer gels, cells, or other substrates. Such supports may take the form, for example, of beads, wells, dipsticks, or membranes.

"Target region" refers to a region of the nucleic acid which is amplified and/or detected. The term "target sequence" refers to a sequence with which a probe or primer will form a stable hybrid under desired conditions.

II. Positional Cloning of Mouse SAC1 Gene and the Discovery of a Gene and Its Sequence Variation Associated With Altered Sensation for Carbohydrates

Inbred strains of mice differ in their intake of sweeteners (Bachmanov A.A., Reed D.R., Tordoff M.G., Price R.A., and Beauchamp G.K. Intake of ethanol, sodium chloride, sucrose, citric acid, and quinine hydrochloride solutions by mice: a genetic analysis. *Behavior Genetics*, 1996;26:563-573; Lush I.E., The genetics of tasting in mice. VI. Saccharin, acesulfame, dulcin and sucrose. *Genet Res*, 1989;53:95-99; Lush I. The genetics of bitterness, sweetness, and saltiness in strains of mice. In *Genetics of Perception and Communication*, Vol. 3, eds. Wysocki C. and Kare M., New York: Marcel Dekker, 1991:227-235; Capretta P.J. Saccharin and saccharin-glucose ingestion in two inbred strains of *Mus musculus*. *Psychon. Sci.*, 1970;21:133-135; Nachman M. The inheritance of saccharin preference. *Journal of Comp Physiol Psychol*, 1959;52:451-457). Breeding and linkage experiments suggest that a single gene, the *Sac* locus (for saccharin intake), accounts for a large proportion of the genetic variance (Fuller J.L. Single-locus control of saccharin preference in mice. *Journal of Heredity*, 1974;65:33-36; Capeless C.G. and Whitney G. The genetic basis of preference for sweet substances among inbred strains of mice: preference ratio phenotypes and the alleles of the Sac and dpa loci. *Chem Senses*, 1995;20:291-298; Bachmanov A.A. et al. Sucrose consumption in mice: major influence of two genetic loci affecting peripheral sensory responses. *Mammalian Genome*, 1997;8:545-548; Belknap J.K. et al. Single-locus control of saccharin intake in BXD/Ty recombinant inbred (RI) mice: some methodological implications for RI strain analysis. *Behav Genet*, 1992;22:81-100; Blizard D.A., Kotlus B., and Frank M.E. Quantitative trait loci associated with short-term intake of sucrose, saccharin and quinine solutions in laboratory mice. *Chem Senses*, 1999;24:373-85). Using genetic and physical mapping methods, an interval of 194 kb was identified at the telomeric end of mouse chromosome 4 that contains the *Sac* locus. BAC sequencing within this interval led to the identification of a gene that has a 30% amino acid homology with other putative taste receptors (Hoon M.A. et al. Putative mammalian taste receptors: a class of taste-specific GPCRs with distinct topographic selectivity. *Cell*, 1999;96:541-551). This gene is expressed in mouse tongue. Mutation detection on this gene revealed a missense mutation (Ile61Thr) with four other sequence variants define a haplotype found in mice

with low sweetener preference (129, Balb/c, AKR, and DBA2). An alternative five variant haplotype is found in mice with a high preference for sweet fluids (B6, SWR, IS, ST, and SEA). A human ortholog of this gene exists, and is expressed in human taste papillae. We therefore suggest that this gene is a sweet taste receptor, and variation within this gene is responsible for the phenotype of the *Sac* locus.

To identify this locus, mice from the high sweetener preference (C57BL/6ByJ; B6) and the low sweetener preference (129P3/J; formerly 129/J, abbreviated here as 129) were used as parental strains to produce an F2 generation. The F2 mice were phenotyped for sweetener preference using 96-hour two-bottle taste tests and genotyped with markers polymorphic between the B6 and 129 strains (Fig. 1A). The results of this analysis indicated peak linkage near marker D18346 with the B6 allele having a dominant mode of inheritance. Using recombinant mice from the F2 generation, 129.B6-*Sac* partially congenic mice were created, using genotypic (B6 allele at D18346; Fig. 1B) and phenotypic (high saccharin intake; Fig. 1C) characteristics as selection criteria for each generation. Genotyping of partially congenic mice with polymorphic markers defined the *Sac* nonrecombinant interval. Radiation hybrid mapping was conducted with additional markers (R74924, D18402, D18346, Agrin, V2r2 and D4Ertd296e). These markers were amplified using DNA and mouse and hamster control DNA in the T31 mouse radiation hybrid panel, scored for the presence or absence of an appropriately sized band, and the data analyzed by the Jackson Laboratory. All markers were within the SAC1 confidence interval suggested by the initial linkage analysis, and were used in subsequent analyses.

A BAC library was screened with markers within the nonrecombinant interval, and a contig was developed (Fig. 1D). A BAC clone was selected for sequencing (RPCI-23-118E21, 246 kb). Within this BAC, a gene with a 30% homology to T1R1 (a putative taste receptor) was discovered (Fig. 2A), along with other ESTs and known genes (Table 1). The human ortholog to this gene was identified from a BAC available in the public *htgs* database, and the predicted protein sequence was aligned with SAC1 and T1R1. SAC1 is 858 amino acids in length and contains six exons; the intron and exon boundaries were determined by

sequencing of the mouse tongue cDNA (Fig. 2B). The secondary structure of this protein with regards to transmembrane domains was predicted (Fig. 2C).

To determine whether this gene might contain functional polymorphisms that could account for the behavioral differences between the two strains, 11.8 kb of sequence, including the SAC1 gene and several kb up and downstream were amplified with PCR primers and then sequenced using DNA from the high and low preferring strains (Lush I.E., The genetics of tasting in mice. VI. Saccharin, acesulfame, dulcin and sucrose. *Genet Res*1989;53:95-99; Lush I. The genetics of bitterness, sweetness, and saltiness in strains of mice. In *Genetics of Perception and Communication*, Vol. 3, eds. Wysocki C. and Kare M., New York: Marcel Dekker, 1991:227-235). Many variants existed between these strains, and of these, five variants were found in the low preferring strains but not in the high preferring strain. One of these variants results in a missense mutation (Ile61Thr; Fig. 2). The other four variants were in non-coding regions (T>A -2383 nt; A>G -183 nt; A>G +134 nt; T>C +651 nt, between exon 2 and 3). These five variants will be referred to as the 129-like or B6-like haplotypes. Additional inbred strains of mice with known saccharin and sucrose preferences (Lush I.E., The genetics of tasting in mice. VI. Saccharin, acesulfame, dulcin and sucrose. *Genet Res*, 1989;53:95-99; Lush I. The genetics of bitterness, sweetness, and saltiness in strains of mice. In *Genetics of Perception and Communication*, Vol. 3, eds. Wysocki C. and Kare M., New York: Marcel Dekker, 1991:227-235; Lush I.E. and Holland G. The genetics of tasting in mice. V. Glycine and cyclohexamide. *Genet Res*, 1988;52:207-212) were also sequenced. The 129-like haplotype was found in mice with lower sweetener preference and the B6-like haplotype was found in mice with higher sweetener preference (Fig. 3).

B6 mice have higher maximal gustatory neural firing in response to sweeteners compared with 129 mice, as do the 129.B6-*Sac* partially congenic strains (Bachmanov A.A. et al. Sucrose consumption in mice: major influence of two genetic loci affecting peripheral sensory responses. *Mammalian Genome*, 1997;8:545-548). Thus, the SAC1 gene is likely to be expressed in tongue. To test this hypothesis, RNA from mouse and human tongue was extracted, reversed transcribed into cDNA and primers, chosen to span an intron, were used in a PCR

reaction. Genomic and cDNA yielded bands of different sizes, which were purified and sequenced (Figure 4AB). Sequencing results confirmed that the bands were derived from this gene with the appropriate intron/exon boundaries. Further analysis of expression in cDNA in mouse tissue, using commercially available mouse cDNA, indicated this gene is also expressed is widely expressed. The broad range of tissue expression of this gene may indicate that other tissues use this receptor to sense extra cellular sugars (Fig. 4A).

Hoon et al. identified a gene, *Gpr70* (formerly TR1 or T1R1) as a putative sweet receptor based mainly on its expression in anterior tongue taste cells. Since it also mapped to distal chromosome 4, it was a logical candidate for SAC1. However, we have shown that *Gpr70* is at least 4 cM proximal to SAC1 (Li X. et al. The saccharin preference locus (*Sac*) and the putative sweet taste receptor (*Gpr70*) gene have distinct locations on mouse chromosome 4. *Mammalian Genome*, 2001;12:13-16). Nevertheless, *Gpr70* could be an additional sweet receptor and there could be others. It has been argued based upon human psychophysical studies and studies of sweet taste transduction mechanisms that there must be more than one sweet receptor. Other lines of evidence, however, are more consistent with the existence of one or a very few receptors (Bartoshuk L.M. Is sweetness unitary? An evaluation of the evidence for multiple sweeteners. In *Sweetness*, ed. Dobbing, J., London: Springer-Verlag, 1987:33-46). At present no evidence has been found of a family of *Sac*-like receptors resembling the large family of bitter receptors recently reported (Matsunami H., Montmayeur J.P., and Buck L.B. A family of candidate taste receptors in human and mouse [see comments]. *Nature*, 2000;404:601-604; Adler E. et al. A novel family of mammalian taste receptors [see comments]. *Cell*, 2000;100:693-702). The sweet substances that exist in nature, which presumably shaped the evolution of sweet receptor(s), are likely much more similar amongst themselves, mostly simple sugars, than are the vast array of structurally diverse bitter tasting compounds.

A receptor for the sugar trehalose has recently been identified in the fruit fly, *Drosophila melanogaster.* Surprisingly, the trehalose and other fly taste receptors, have no homology with SAC1. The specialization of flies for the sugar trehalose may account for this divergence.

There may be multiple sweet receptors; evidence from across species comparisons, psychophysical cross adaptation, and sweetness competitors has been reviewed (Bartoshuk L.M. Is sweetness unitary? An evaluation of the evidence for multiple sweeteners. In *Sweetness*, ed. Dobbing, J., London: Springer-Verlag, 1987:33-46). The SAC1 gene accounts for ~40% of the genetic differences in sweet perception between these two particular strains of mice, but other receptors, and other alleles of these receptors may exist.

Because sucrose is perceived to be bad for human health, considerable resources are directed toward the discovery of high potency, low caloric sweeteners. Most of the most widely known high potency sweeteners were discovered serendipitously, i.e., the sweetener was synthesized for a different purpose and someone in the laboratory accidentally tasted it and discovered it was sweet (Walters E.D. The rational discovery of sweeteners. In *Sweeteners. Discovery, molecular design, and chemoreception*, eds. Walters D.E., Orthoefer F.T., and DuBois G.E., American Chemical Society, USA, 1991:1-11). More direct methods, however, have been employed to identify new sweet compounds, and the sweet receptor has been extensively modeled to predict which ligands will be sweet.

It is not known how or why different alleles of SAC1 arose in inbred strains of mice but their existence, in addition to providing us with a tool to identify a sweet receptor, raises the question of whether they might also characterize human populations. There appear to exist reliable individual differences in human sensitivity and preference for sweet sugars but whether these are genetically influenced remains to be determined. The identification of SAC1 should facilitate research in this area. Also, the observation that SAC1 is expressed in several tissues in addition to tongue raises the interesting possibility that it could be involved in other aspects of sugar recognition and that allelic variants in this gene could be related to diseases or conditions such as diabetes and obesity.

Alleles of the gene described in this application are likely to account for the SAC1 behavioral and neurological phenotype for four reasons. First, the SAC1 nonrecombinant region is small, less than 194 kb; this gene lies within this

nonrecombinant interval and the peak of LOD score corresponds closely with the location of the gene. Second, of the genes contained within this region, no others are viable candidates for SAC1. Third, this gene has sequence homology to other putative taste receptors, and is expressed in the tongue. Finally, a haplotype with a missense mutation is found in mice with low sweetener preference but not in mice with high sweetener preference. These data strongly suggest that mutations of this gene account for differences in the acceptance and preference for sweeteners attributed to the SAC1 locus.

Among the multiple mechanisms involved in regulation of ethanol intake, one of the least appreciated factors is the perception of its flavor (Nachman M., Larue C., Le Magnen J. The role of olfactory and orosensory factors in the alcohol preference of inbred strains of mice. *Physiology Behavior*, 1971;6:53-95). Although individual variability in the perception of ethanol flavor by adults and children was described over 60 years ago (Richter C.P. Alcohol as a food. *Quart. J. Studies Alcohol*, 1941;1:650-62), the hypothesis that individual differences in alcohol chemosensory perception can affect alcohol intake did not receive due attention. As a result, the relationship between alcohol chemosensation and intake is not well-understood. Humans perceive ethanol flavor as a combination of components, including sweetness, bitterness, odor and irritation (burning sensation), which depend on ethanol concentration (Green B.G. The sensitivity of the tongue to ethanol. *Ann. NY. Acad. Sci.*, 1987;510:315-7; Bartoshuk L.M., Conner E., Grubin D., Karrer T., Kochenbach K., Palsco M., et al. PROP supertasters and the perception of ethyl alcohol. *Chem. Senses*, 1993.). Rats detect sweet (sucrose-like) and bitter (quinine-like) sensory components in ethanol (Kiefer S.W., Lawrence G.J. The sweet-bitter taste of alcohol: aversion generalization to various sweet-quinine mixtures in the rat. *Chem. Senses*, 1988;13:633-41; Kiefer S.W., Mahadevan R.S. The taste of alcohol for rats as revealed by aversion generalization tests. *Chem. Senses*, 1993;18:509-22) and probably perceive the other components detected by humans as well.

The relationship between ethanol and sweetener perception and consumption has been studied the most and is supported by several lines of evidence:

(a) Electrophysiological recordings from gustatory nerves indicate that lingual application of ethanol activates sweetener-responsive neural fibers (Hellekant G., Danilova V., Roberts T., Ninomiya Y. The taste of ethanol in a primate model: I. Chorda tympani nerve response in Macaca mulatta. *Alcohol*, 1997;14:473-84; Sako N., Yamamoto T. Electrophysiological and behavioral studies on taste effectiveness of alcohols in rats. *Am. J. Physiol.*, 1999;276:R388-96).

(b) Conditioned taste aversions generalize between ethanol and sucrose (Kiefer S.W., Lawrence G.J. The sweet-bitter taste of alcohol: aversion generalization to various sweet-quinine mixtures in the rat. *Chem. Senses*, 1988;13:633-41; Kiefer S.W., Mahadevan R.S. The taste of alcohol for rats as revealed by aversion generalization tests. *Chem. Senses*, 1993;18:509-22; Lawrence G.J., Kiefer S.W. Generalization of specific taste aversions to alcohol in the rat. *Chem. Senses*, 1987;12:591-9; Blizard D.A., McClearn G.E. Association between ethanol and sucrose intake in the laboratory mouse: exploration via congenic strains and conditioned taste aversion. *Alcohol. Clin. Exp. Res.*, 2000;24:253-8.), suggesting that ethanol and sucrose share the same taste property, most likely sweetness.

(c) Genetic associations between preferences for ethanol and sweeteners were found among some rat and mouse strains and within their segregating crosses (Overstreet D.H., Kampov-Polevoy A.B., Rezvani A.H., Murelle L., Halikas J.A., Janowsky D.S. Saccharin intake predicts ethanol intake in genetically heterogeneous rats as well as different rat strains. *Alcohol. Clin. Exp. Res.*, 1993;17:366-9; Sinclair J.D., Kampov-Polevoy A., Stewart R., Li T-K. Taste preferences in rat lines selected for low and high alcohol consumption. *Alcohol*, 1992;9:155-60; Stewart R.B., Russell R.N., Lumeng L., Li T-K., Murphy J.M. Consumptions of sweet, salty, sour, and bitter solutions by selectively bred alcohol-preferring and alcohol-nonpreferring lines of rats. *Alcohol. Clin. Exp. Res.*, 1994;18:375-81; Belknap J.K., Crabbe J.C., Young E.R. Voluntary consumption of alcohol in 15 inbred mouse strains. *Psychopharmacol.*, 1993;112:503-10;

Bachmanov A.A., Reed D.R., Tordoff M.G., Price R.A., Beauchamp G.K. Intake of ethanol, sodium chloride, sucrose, citric acid, and quinine hydrochloride solutions by mice: a genetic analysis. *Behav. Genet.*, 1996;26:563-73; Bachmanov A.A., Tordoff M.G., Beauchamp G.K. Ethanol consumption and taste preferences in C57BL/6ByJ and 129/J mice. *Alcohol. Clin. Exp. Res.*, 1996;20:201-6), reviewed in (Kampov-Polevoy A.B., Garbutt J.C., Janowsky D.S. Association between preference for sweets and excessive alcohol intake: a review of animal and human studies. *Alcohol. Alcohol.*, 1999;34:386-95; Overstreet D.H., Rezvani A.H., Parsian A. Behavioural features of alcohol-preferring rats: focus on inbred strains. *Alcohol. Alcohol.*, 1999;34:378-85); with some exceptions (Phillips T.J., Crabbe J.C., Metten P., Belknap J.K. Localization of genes affecting alcohol drinking in mice. *Alcohol. Clin. Exp. Res.*, 1994;18:931-941; Parsian A., Overstreet D.H., Rezvani A.H. Independent segregation of alcohol and saccharin intakes in F2 progeny from FH/ACI intercross (Abstract). *Alcohol. Clin. Exp. Res.*, 2000;24(Supplement):58A)).

(d) Human studies show that alcoholics have a stronger liking of concentrated sucrose compared with nonalcoholics (Kampov-Polevoy A.B., Garbutt J.C., Davis C.E., Janowsky D.S. Preference for higher sugar concentrations and Tridimensional Personality Questionnaire scores in alcoholic and nonalcoholic men. *Alcohol. Clin. Exp. Res.*, 1998;22:610-4; Kampov-Polevoy A.B., Garbutt J.C., Janowsky D. Evidence of preference for a higher concentration sucrose solution in alcoholic men. *American Journal of Psychiatry*, 1997;154:269-70).

There are several possible mechanisms that could underlie the association between sweetener and ethanol responses:

(a) *Common peripheral taste mechanisms*, which may involve the interaction of ethanol with a peripheral sweet taste transduction. At least one such common peripheral mechanism is mediated by the *Gpr98* gene (SAC1 locus) encoding a sweet taste receptor (as described below).

(b) *Common brain mechanisms*. The regulation of ingestive responses to ethanol and sweeteners may involve common opioidergic, serotonergic and dopaminergic brain neurotransmitter systems (Gosnell B.A., Majchrzak M.J. Centrally administered opioid peptides stimulate saccharin intake in nondeprived rats. *Pharm. Biochem. Behav.*, 1989;33:805-10; George S.R., Roldan L., Lui A., Naranjo C.A. Endogenous opioids are involved in the genetically determined high preference for ethanol consumption. *Alcohol. Clin. Exp. Res.*, 1991;15:668-72; Hubell C.L., Marglin S.H., Spitalnic S.J., Abelson M.L., Wild K.D., Reid L.D. Opioidergic, serotoninergic, and dopaminergic manipulations and rats' intake of a sweetened alcoholic beverage. *Alcohol*, 1991;8:355-67; Pucilowski O., Rezvani A.H., Janowsky D.S. Suppression of alcohol and saccharin preference in rats by a novel Ca^{2+} channel inhibitor, Goe 5438. *Psychopharmacol.*, 1992;107:447-52). These mechanisms could be responsible for the emotional response to the pleasantness of ethanol or sweeteners, or the motivational mechanisms driving their intakes.

(c) *Common signals related to the caloric value of ethanol and sugars* (Gentry R.T., Dole V.P. Why does a sucrose choice reduce the consumption of alcohol in C57BL/6J mice? *Life Sci.*, 1987;40:2191-4). Ethanol is metabolized in the body through some of the same pathways as carbohydrates and provides comparable energy. Thus, energy derived from carbohydrates and ethanol may have similar rewarding effects through the same hunger and satiety mechanisms.

(d) *Incidental genetic linkage*. Different genes affecting responses to ethanol and sweeteners may reside nearby on the same chromosome.

Ethanol consumption is a complex trait, depending on multiple mechanisms of its regulation and determined by multiple genes. A body of evidence suggests that ethanol consumption may depend on perception of its flavor, and that there is an association between perception and consumption of ethanol and sweet-tasting compounds. However, only a few genes have been identified as candidates affecting ethanol consumption.

The present invention provides that a gene, SAC1, is associated with the detection of a sensing of carbohydrates, other sweet compounds, and alcohols including ethanol. The sequence of the mouse SAC1 cDNA (SEQ ID NO: 1) is:

ATGCCAGCTTTGGCTATCATGGGTCTCAGCCTGGCTGCTTTCCTGGAGC
TTGGGATGGGGGCCTCTTTGTGTCTGTCACAGCAATTCAAGGCACAAG
GGGACTACATACTGGGCGGGCTATTTCCCCTGGGCTCAACCGAGGAGG
CCACTCTCAACCAGAGAACACAACCCAACAGCATCCCGTGCAACAGGT
TCTCACCCCTTGGTTTGTTCCTGGCCATGGCTATGAAGATGGCTGTGGA
GGAGATCAACAATGGATCTGCCTTGCTCCCTGGGCTGCGGCTGGGCTA
TGACCTATTTGACACATGCTCCGAGCCAGTGGTCACCATGAAATCCAG
TCTCATGTTCCTGGCCAAGGTGGGCAGTCAAAGCATTGCTGCCTACTG
CAACTACACACAGTACCAACCCCGTGTGCTGGCTGTCATCGGCCCCCA
CTCATCAGAGCTTGCCCTCATTACAGGCAAGTTCTTCAGCTTCTTCCTC
ATGCCACAGGTCAGCTATAGTGCCAGCATGGATCGGCTAAGTGACCGG
GAAACGTTTCCATCCTTCTTCCGCACAGTGCCCAGTGACCGGGTGCAG
CTGCAGGCAGTTGTGACTCTGTTGCAGAACTTCAGCTGGAACTGGGTG
GCCGCCTTAGGGAGTGATGATGACTATGGCCGGGAAGGTCTGAGCATC
TTTTCTAGTCTGGCCAATGCACGAGGTATCTGCATCGCACATGAGGGC
CTGGTGCCACAACATGACACTAGTGGCCAACAGTTGGGCAAGGTGCTG
GATGTACTACGCCAAGTGAACCAAAGTAAAGTACAAGTGGTGGTGCTG
TTTGCCTCTGCCCGTGCTGTCTACTCCCTTTTTAGTTACAGCATCCATCA
TGGCCTCTCACCCAAGGTATGGGTGGCCAGTGAGTCTTGGCTGACATC
TGACCTGGTCATGACACTTCCCAATATTGCCCGTGTGGGCACTGTGCTT
GGGTTTTTGCAGCGGGGTGCCCTACTGCCTGAATTTTCCCATTATGTGG
AGACTCACCTTGCCCTGGCCGCTGACCCAGCATTCTGTGCCTCACTGAA
TGCGGAGTTGGATCTGGAGGAACATGTGATGGGGCAACGCTGTCCACG
GTGTGACGACATCATGCTGCAGAACCTATCATCTGGGCTGTTGCAGAA
CCTATCAGCTGGGCAATTGCACCACCAAATATTTGCAACCTATGCAGC
TGTGTACAGTGTGGCTCAAGCCCTTCACAACACCCTACAGTGCAATGT
CTCACATTGCCACGTATCAGAACATGTTCTACCCTGGCAGCTCCTGGA
GAACATGTACAATATGAGTTTCCATGCTCGAGACTTGACACTACAGTT

TGATGCTGAAGGGAATGTAGACATGGAATATGACCTGAAGATGTGGGT
GTGGCAGAGCCCTACACCTGTATTACATACTGTGGGCACCTTCAACGG
CACCCTTCAGCTGCAGCAGTCTAAAATGTACTGGCCAGGCAACCAGGT
GCCAGTCTCCCAGTGTTCCCGCCAGTGCAAAGATGGCCAGGTTCGCCG
AGTAAAGGGCTTTCATTCCTGCTGCTATGACTGCGTGGACTGCAAGGC
GGGCAGCTACCGGAAGCATCCAGATGACTTCACCTGTACTCCATGTAA
CCAGGACCAGTGGTCCCCAGAGAAAAGCACAGCCTGCTTACCTCGCAG
GCCCAAGTTTCTGGCTTGGGGGGAGCCAGTTGTGCTGTCACTCCTCCTG
CTGCTTTGCCTGGTGCTGGGTCTAGCACTGGCTGCTCTGGGGCTCTCTG
TCCACCACTGGGACAGCCCTCTTGTCCAGGCCTCAGGTGGCTCACAGT
TCTGCTTTGGCCTGATCTGCCTAGGCCTCTTCTGCCTCAGTGTCCTTCTG
TTCCCAGGGCGGCCAAGCTCTGCCAGCTGCCTTGCACAACAACCAATG
GCTCACCTCCCTCTCACAGGCTGCCTGAGCACACTCTTCCTGCAAGCAG
CTGAGACCTTTGTGGAGTCTGAGCTGCCACTGAGCTGGGCAAACTGGC
TATGCAGCTACCTTCGGGGACTCTGGGCCTGGCTAGTGGTACTGTTGG
CCACTTTTGTGGAGGCAGCACTATGTGCCTGGTATTTGATCGCTTTCCC
ACCAGAGGTGGTGACAGACTGGTCAGTGCTGCCCACAGAGGTACTGG
AGCACTGCCACGTGCGTTCCTGGGTCAGCCTGGGCTTGGTGCACATCA
CCAATGCAATGTTAGCTTTCCTCTGCTTTCTGGGCACTTTCCTGGTACA
GAGCCAGCCTGGCCGCTACAACCGTGCCCGTGGTCTCACCTTCGCCAT
GCTAGCTTATTTCATCACCTGGGTCTCTTTTGTGCCCCTCCTGGCCAAT
GTGCAGGTGGCCTACCAGCCAGCTGTGCAGATGGGTGCTATCCTAGTC
TGTGCCCTGGGCATCCTGGTCACCTTCCACCTGCCCAAGTGCTATGTGC
TTCTTTGGCTGCCAAAGCTCAACACCCAGGAGTTCTTCCTGGGAAGGA
ATGCCAAGAAAGCAGCAGATGAGAACAGTGGCGGTGGTGAGGCAGCT
CAGGGACACAATGAATGA

The geonomic DNA sequence of the mouse SAC1 gene (SEQ ID NO: 2) is:

ATCTGAGCCTTAGACACAGCACTGGTGCCAGGCAAACACTCCTGGGCC
TACATGCTTGGG

GCCTCTTCATATTCCAAAAGCTGTCTTTGGGTAAGATGAAGTTCCTCTG
GCAGTGGCATG
AGTGCTGAAGGCTCTTTCCCTGCCCTTCACCTGCTTTCTTGATAGTCTCT
CTGCATACCA
AACAGGCCCTTGTCTCCTGGGAAATGGAAACTATGAAATCAATAGCTG
AGGCTTCTCTAG
GAAAGCCTGCCCTGGTCAGTACAACCTGTTTCACAGCTTCTATAGAAT
AGTTACATCAGC
CTTCTGAAGATGGCCTCTTAGAGCACATGCACCCCCAAGATTCTAAGA
TGTCAATACTAA
CTGACCAAACCATACCTCTCTAGCCAGCCCTGCTGCTCCTGTTGTCTGG
TACCCAGGTGA
CTGAGGACATGACTGGTGGAAGGAAACTAGGCCCCTTTGTCTGTCAGA
TGGCCATACCCA
GCATGGCTGATGCCCAGTGTATAAGACCCTACGCTTTTCCACTGGTCTT
AATGTTAAACC
CTAGGACAGTGTCCTCAGCATAGCTGGTGTGTGTGAATGCAAACTTTG
GGGCATATCTCT
TCCATTAAGCACTGTGATATATGTAGTATTTCCAACAAATAAATTATAC
CTACATGATTG
GGTATAGCATTCTGGGATGGGTCACAGGTGTGTCAGGTGCCTAATTAT
GTGGGGGAAGAA
CATAGAAATATATAGGTGGGGAGGGAGCTAACCCTAGGAATAAGGCT
AAAGCATGTGTCT
CCAGTCCTGAAGACTCAAAGGGCAACGTGAATCATGAGACATGTTCAG
GACTGAAGGAGT
TGCCATGTATCTGTCCTTGATGTATCTTAATCATACATACACTATGAGA
TCTGTGTTACC
TCCATTTTGCAGGTGAGAAAAGAAACACCTGAATGGCCTACCTTAAAG
GGCTAAGTGGGA
AAATAGGTCTGAAGATAACCCAGGCACTGTGTGACAAAGCGGGAAGA
AAACTAGAGATGC

TTTCTTCATGGCAACAACCTAGAGGGTACAACCTAGTGGTTTCTTCTTG
GTACTCCACTG
TATACACCCCATCTGCTTGGGCTGTACATTGTCTGACCATGCTTATAAC
AAAAGTCACAT
ACTACTAGCCAAGACTGAGAACTTAGAGCGACTGGCCAGAAAGTAAA
GATACAACAGTTG
ATATGTGTGCCACACACAGATCCATGTGTACATGTCTATTAATTATGTG
AACGTGCTTTG
TGGACATCCTCACAAAGCAGCAGGGAAATGCAAAGGTCATTTCCATAA
CACCTGCTGGAC
ACCATATGACATTGAGATTACCGGGGTGCCCATTCCAACAAGAGTTAA
TAGCTCCCCCTA
TGTTTGGGTGCCAGAAACCTGATTTGTTAGCAATAGCTCCCTCACATCC
AGATTAAGAGG
GGGATGGCTTAGCTAGGGTTACTATGATGAAACTATGACCAAAGCAAC
TTGTGGGTAAAA
GGGTGTATTTGGCTTACACTTCCATATCACTTCATCAAAGTGAGGACA
GGAACTCAAATA
GAGTAGGAATTTGGTGACAAGAGCTGATGTAGAGGCAATGCAGTGGT
GCCACTTAGTGGC
GCGCTCAGTCTGCTCCCTTTCTTAATAGAATGCAAGACCACCAGCCCAT
GGGTGGCACCA
CAATGGGACCGGGCCCTTCCCCATCGGTCACTAAGAAAATGCCCTACA
GCCAGATCTTAT
GGAGACATTTTCTCAACGGAGGCTCACTCCTTTCAGATAACTCTATATC
AAATTGACATA
AACCAGAACAGAGGAGGAGGCTAAGAAGGAAACTGCCAATTGCATAC
ATGCACACACCTG
GCCCTAGCAGCTGCAGGAAGCTATTTGTTTATGGCCTTTTCTCATTTTC
ATGGACCAGCA
TGAGCACTCTGCAGAGAGAGATGCCTGCATGCCTGCCAAGGCAGGAGT
GCTTACACTGAA

GGTCAACAGGATGGCAGGGGGGCTGCAGAGCTTCCAAGTGTCAGAAC
CCCAGCAGAAGAG
CTGAGACCCTTGCCCGAGGACTCAGGCGGGTTGGGAAGGCCAGGAAA
TTCAGCCAGAGCT
CTTCTTCAGATGGGGTACCATCTGAAGGTTAGACCAGCTAGCCAGCTG
TTGTTGAGGGAC
CACCTCTGCAGCCCCTACCTTTGGAAGATAGAAAGTGTCTCTGTGACA
AGTATGGCCATT
GTGCCCCCTTATTCCACAGTCAACAGAAACCCTGGAATCCTGAACACT
TCTGCAGCTTCT
TTTTTACAGTCTGCCAGGTTGCTCTAGGAATGAAGGGTGCCGAGAGGC
TTGGGCGTAGGC
AGGTGACAAGACCACAGTTAGTGGTCACAGCTGGCTTACTGGATCACT
CTTGGACAGAGT
TTGTTAGATATGGAGTGGAGTATACACAAGGCATCAGGCGGGGGATAT
TGAATGTATCAC
CGGAGCTCCTTGGGGCTTGGCAGCCAAGCACAGCAGTGGTTTTGCTAA
ACAAATCCACGG
TTCCCTCCCCTTGACGCAGTACATCTGTGGCTCCAACCCCACACACCCA
CCCATTGTTAG
TGCTGGAGACTTCTACCTACCATGCCAGCTTTGGCTATCATGGGTCTCA
GCCTGGCTGCT
TTCCTGGAGCTTGGGATGGGGGCCTCTTTGTGTCTGTCACAGCAATTCA
AGGCACAAGGG
GACTACATACTGGGCGGGCTATTTCCCCTGGGCTCAACCGAGGAGGCC
ACTCTCAACCAG
AGAACACAACCCAACAGCATCCCGTGCAACAGGTATGGAGGCTAGTA
GCTGGGGTGGGAG
TGAACCGAAGCTTGGCAGCTTTGGCTCCGTGGTACTACCAATCTGGGA
AGAGGTGGTGAT
CAGTTTCCATGTGGCCTCAGGTTCTCACCCCTTGGTTTGTTCCTGGCCA
TGGCTATGAAG

ATGGCTGTGGAGGAGATCAACAATGGATCTGCCTTGCTCCCTGGGCTG
CGGCTGGGCTAT
GACCTATTTGACACATGCTCCGAGCCAGTGGTCACCATGAAATCCAGT
CTCATGTTCCTG
GCCAAGGTGGGCAGTCAAAGCATTGCTGCCTACTGCAACTACACACAG
TACCAACCCCGT
GTGCTGGCTGTCATCGGCCCCCACTCATCAGAGCTTGCCCTCATTACAG
GCAAGTTCTTC
AGCTTCTTCCTCATGCCACAGGTGAGCCCACTTCCTTTGTGTTCTCAAC
CGATTGCACCC
ATTGAGCTCTCATATCAGAAAGTGCTTCTTGATCACCACAGGTCAGCT
ATAGTGCCAGCA
TGGATCGGCTAAGTGACCGGGAAACGTTTCCATCCTTCTTCCGCACAG
TGCCCAGTGACC
GGGTGCAGCTGCAGGCAGTTGTGACTCTGTTGCAGAACTTCAGCTGGA
ACTGGGTGGCCG
CCTTAGGGAGTGATGATGACTATGGCCGGGAAGGTCTGAGCATCTTTT
CTAGTCTGGCCA
ATGCACGAGGTATCTGCATCGCACATGAGGGCCTGGTGCCACAACATG
ACACTAGTGGCC
AACAGTTGGGCAAGGTGCTGGATGTACTACGCCAAGTGAACCAAAGT
AAAGTACAAGTGG
TGGTGCTGTTTGCCTCTGCCCGTGCTGTCTACTCCCTTTTTAGTTACAGC
ATCCATCATG
GCCTCTCACCCAAGGTATGGGTGGCCAGTGAGTCTTGGCTGACATCTG
ACCTGGTCATGA
CACTTCCCAATATTGCCCGTGTGGGCACTGTGCTTGGGTTTTTGCAGCG
GGGTGCCCTAC
TGCCTGAATTTTCCCATTATGTGGAGACTCACCTTGCCCTGGCCGCTGA
CCCAGCATTCT
GTGCCTCACTGAATGCGGAGTTGGATCTGGAGGAACATGTGATGGGGC
AACGCTGTCCAC

GGTGTGACGACATCATGCTGCAGAACCTATCATCTGGGCTGTTGCAGA
ACCTATCAGCTG
GGCAATTGCACCACCAAATATTTGCAACCTATGCAGCTGTGTACAGTG
TGGCTCAAGCCC
TTCACAACACCCTACAGTGCAATGTCTCACATTGCCACGTATCAGAAC
ATGTTCTACCCT
GGCAGGTAAGGGTAGGGTTTTTTGCTGGGTTTTGCCTGCTCCTGCAGG
AACACTGAACCA
GGCAGAGCCAAATCTTGTTGTGACTGGAGAGGCCTTACCCTGACTCCA
CTCCACAGCTCC
TGGAGAACATGTACAATATGAGTTTCCATGCTCGAGACTTGACACTAC
AGTTTGATGCTG
AAGGGAATGTAGACATGGAATATGACCTGAAGATGTGGGTGTGGCAG
AGCCCTACACCTG
TATTACATACTGTGGGCACCTTCAACGGCACCCTTCAGCTGCAGCAGT
CTAAAATGTACT
GGCCAGGCAACCAGGTAAGGACAAGACAGGCAAAAAGGATGGTGGGT
AGAAGCTTGTCGG
TCTTGGGCCAGTGCTAGCCAAGGGGAGGCCTAACCCAAGGCTCCATGT
ACAGGTGCCAGT
CTCCCAGTGTTCCCGCCAGTGCAAAGATGGCCAGGTTCGCCGAGTAAA
GGGCTTTCATTC
CTGCTGCTATGACTGCGTGGACTGCAAGGCGGGCAGCTACCGGAAGCA
TCCAGGTGAACC
GTCTTCCCTAGACAGTCTGCACAGCCGGGCTAGGGGGCAGAAGCATTC
AAGTCTGGCAAG
CGCCCTCCCGCGGGGCTAATGTGGAGACAGTTACTGTGGGGGCTGGCT
GGGGAGGTCGGT
CTCCCATCAGCAGACCCCACATTACTTTTCTTCCTTCCATCACTACAGA
TGACTTCACCT
GTACTCCATGTAACCAGGACCAGTGGTCCCCAGAGAAAAGCACAGCCT
GCTTACCTCGCA

GGCCCAAGTTTCTGGCTTGGGGGGAGCCAGTTGTGCTGTCACTCCTCCT
GCTGCTTTGCC
TGGTGCTGGGTCTAGCACTGGCTGCTCTGGGGCTCTCTGTCCACCACTG
GGACAGCCCTC
TTGTCCAGGCCTCAGGTGGCTCACAGTTCTGCTTTGGCCTGATCTGCCT
AGGCCTCTTCT
GCCTCAGTGTCCTTCTGTTCCCAGGGCGGCCAAGCTCTGCCAGCTGCCT
TGCACAACAAC
CAATGGCTCACCTCCCTCTCACAGGCTGCCTGAGCACACTCTTCCTGCA
AGCAGCTGAGA
CCTTTGTGGAGTCTGAGCTGCCACTGAGCTGGGCAAACTGGCTATGCA
GCTACCTTCGGG
GACTCTGGGCCTGGCTAGTGGTACTGTTGGCCACTTTTGTGGAGGCAG
CACTATGTGCCT
GGTATTTGATCGCTTTCCCACCAGAGGTGGTGACAGACTGGTCAGTGC
TGCCCACAGAGG
TACTGGAGCACTGCCACGTGCGTTCCTGGGTCAGCCTGGGCTTGGTGC
ACATCACCAATG
CAATGTTAGCTTTCCTCTGCTTTCTGGGCACTTTCCTGGTACAGAGCCA
GCCTGGCCGCT
ACAACCGTGCCCGTGGTCTCACCTTCGCCATGCTAGCTTATTTCATCAC
CTGGGTCTCTT
TTGTGCCCCTCCTGGCCAATGTGCAGGTGGCCTACCAGCCAGCTGTGC
AGATGGGTGCTA
TCCTAGTCTGTGCCCTGGGCATCCTGGTCACCTTCCACCTGCCCAAGTG
CTATGTGCTTC
TTTGGCTGCCAAAGCTCAACACCCAGGAGTTCTTCCTGGGAAGGAATG
CCAAGAAAGCAG
CAGATGAGAACAGTGGCGGTGGTGAGGCAGCTCAGGGACACAATGAA
TGACCACTGACCC
GTGACCTTCCCTTTAGGGAACCTAGCCCTACCAGAAATCTCCTAAGCC
AACAAGCCCCGA

ATAGTACCTCAGCCTGAGACGTGAGACACTTAACTATAGACTTGGACT
CCACTGACCTTA
GCCTCACAGTGACCCCTTCCCCAAACCCCCAAGGCCTGCAGTGCACAA
GATGGACCCTAT
GAGCCCACCTATCCTTTCAAAGCAAGATTATCCTTGATCCTATTATGCC
CACCTAAGGCC
TGCCCAGGTGACCCACAAAAGGTTCTTTGGGACTTCATAGCCATACTTT
GAATTCAGAAA
TTCCCCAGGCAGACCATGGGAGACCAGAAGGTACTGCTTGCCTGAACA
TGCCCAGCCCTG
AGCCCTCACTCAGCACCCTGTCCAGGCGTCCCAGGAATAGAAGGCTGG
GCATGTATGTGT
GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTATGTACGTATG
TATGTATGTAT
CAGGACAGAACAAGAAAGACATCAGGCAGAGGACACTCAGGAGGTAG
GCAACATCCAGCC
TTCTCCATCCCTAGCTGAGCCCTAGCCTGTAGGAGAGAACCAGGTCGC
CGCCAGCACCTT
GGACAGATCACACACAGGGTGCGGGTCAGCACCACGGCCAGCGCCAG
CCACGCGGGACCC
CTGGAATCAGCTTCTAGTACCAAGGACAGAAAAGTTGCCGCAAGGCCC
CTTACTGGCCAG
CACCAGGGACAGAGCCACATGCCTAAGCGGCAAGGGACAAGAGCATC
GTCCATCTGCAGG
CAGGATCAGACCCGGGTCAGTTCTGGACTGGCCCCCACACCTGAATCC
CGGAGCAGCTCA
GCTGGAGAAAAGAGAAACAAGCCACACATCAGTCCCATAAAATTAAA
CGCTTTTTTTAGT
GTTTAAAATAGCATTTACACAGAAGCAGCATTTACACAGAAGCAGCTC
TATGTCAACTAC
CCAGTCACTCAGACTTTGACACAGTGTCTAGTGTAGATGTGTGGGGCC
GCTGTGCCGGGA

TGGCAGTGGCACATGATGATGGGCAGCCACCAGAACAGAAACAGAAC
AGGGCCCAGCTCT
GCAGCTCTTGTGTTCACTGTCACCCACCACTGAGACTGAGACAGTGGC
TAGGTGCCAGGT
CTCTCTCCTGTCTCTCCTACTAGCTACCCTTCACATACCTTCAGTACAA
ACTGTGTTGTC
ATGTGCCAAGTAGCAGGTGGGGAAAGGGGCATGCAAACTGCCCCTTTG
GGTAACTAGCTG
CCACCCTTAGAGCAGGCAGGCTAGCAATAAATAAATAAGTTAGACCCC
ACCTGGGCAGCC
AGAGAGGTTTGAAGGCTCTGTCTAACCCCTCAAAAATCCCACCTTGGC
CTGACAGGTGAG
GCCCATGAACTTAGCGACAGTCAGCCTGTGTCCCTGTGCACAGTTCTGT
GAGGCTTTGGG
GCAAGGGGTACCAAGAGCCCAAGAGAGCCTTTCTTGTTCTAAATGGAG
GTCACTTCCAAA
GAAGGGAACCAGGAGGTGGTCCCTGAGACTTGTGCTGAGGACTTAAA
GTCAGAGATGTCT
CCTTACAAGACTCTATAGATACTTGAGCTGTACCACCATCAGCAGCCC
CAAGAGCAGACA
AAATGTCAAGCCAATATCCTGGTGGTATGGCTGCCCTCAGGCCCTCCT
CTGTAGCCTGCT
CCCTCTGCCCTGGCCCAGAGCCCACAGCTGATCTATCCTGGCTGGCCA
CCACCACGGCCA
GCGCAGAGCTCCTGGCACAGCAGGAGCACAGACTCAGCCACAGGCAG
CGCTGAAGACATT
GGTTGATCATCACATGATGTCCACAAAGAACTCACAGGGGTTTCCCAT
GGCCTTTTGGAA
GGACTGGCGGCTACCTGTAAGTTCTGGAGGGACAGCAGCCAGCTCCCG
GACGGGTGGCCC
TCCAGGTGGCCCACCCACTACTGCATAGGCCTTTGTAAGGGGGTGCAG
TGGGGGGAGCCC

TGGGGCAACAGCTGAAGCCTGACTTCGAGGGCTACTGCCACGGCTAAG
CTGGCTGACAGG
CCGCTCCCACCAGCCGGTGCTACCAGACCCACTTGGTACTGTGTGGTCT
GATTCACTGCC
ACTACCCCCAGCTCCAGTTGCCCGGCGCTCCTCTCGGCCTGGGGTCCG
ATGGCTGCTCCG
TGTGGACCCACTGCTCTTGCTCCCTAGGGGGAGGGAAGGGGACAACAG
AGTCAGCACGAG
GCCTGGCCACTTCCAGGGCCACCAGCTGCTCCCAGACAGTCAGGGCAG
GACCTGGTAAGC
CTGGAGATGGTAGGGGAATGGCAGCCATGCAGATACCAGGAACAGCT
GAGAGGCGAGAAG
CTAGGGGCAGTGGCAGACAGCAGGGACAACAGGGGCCAGCCTGGCAC
CCCACACCTAACC
CCAATGCTTGAACCAAGGGTTAATGTTACAGCTGAGAAACTAAAAACC
AGCGAAGGCCCT
GTGTGCCCAGCATTCCCATTAGCCATCCTGGGTTCACCACCCAAAGAC
CCAACCAGGGTC
CACCCAACCCCAGGACCCTGGTCATCTAATTTGCTTAGCCCCTGTCCTG
AAAGTAGTGGG
AACCTGAAAACACGTGCTGGCTGGGGACATGCTGAGAGGGACACAGG
GGGACCTGGCTTA
CCGGCCCGAGAGTCCACTCTGCTAGTCCTTCAGTCTAAGGCTTGCTCAG
CACAAAGCAAG
GGATAGCACAAGTCACACACCAGTCCAGTGCTCACCAATGGCTAATAG
GACGATTTTGGG
CCAAGCTGAGCCTGGGTACATGCAAGGGCCTGTCCATGGTCAGGATTC
ACTCGATAGCTT
CCCCTTGGGCTTTGCCACCCTCTGGCCCAACCTCTCCTGAGTCTTTCTCT
GGACCTTGTA
GCACAAGTGTGCCCCACTCTGCCTAAGACCTCCACATCAGTCCATCTCC
TCCTGAGGGAC

ACCCACCCTTCAAGATCTTCAATATCCCTGGGATATGCTTTAACACTGA
TATGCTTTAAC
AGTGTTGCTTGATACTCTTATCTGGCACTCTGTTGGGATGCAGGCTCCA
TAACTGATAAA
GCCCATTCTCCCCCTAGCTTGGGGCCTAGAGAGTGCCCCTACCTGCTAT
CAGTGGTTACT
TTCATTCTTGCCATATCATCTCCTGGCCTCTTGCCTCTGCCACCTAGCAC
ACCAGGCTGT
CTTCCTATTCTCTAACGGCTTCTACCCACATCAGCCCCTCCCTGTCCCA
CACACTGACTC
TTGAGATGGAACCCACCGGGACTCAAACACACAGCAGGAGCACAGAG
GGAAGCGTCGGGG
CCAGGCAGAGCGTGGGAGTGGGAGGGAGTGGGAGGAGGGGTGGCAC
GCCTCTCACCTTCA
CTCTGCTGGCTCCCAGCACTGCCGCTGCCGCAGCTGAAGCCAGGGTCC
TGGTAAGCAGGC
GGGAAGCAGGGCGGGGGTCCTGGGTACTGGTAGGGGTAGCCTTGACC
CAAGGGCCAGGGT
ACTGATGGGTGGGGCAGTGGGGCCAGTGTGTCCTGATCTGAGGCTCCA
CTGGAGCCACTG
TTGAGGTTCAGGGATGCGAGGTCTGGCAGGGAGGGAGGGAGGGAGGG
GTAAGTGAAGGCA
AATGAATGAGGCCACAGCAACCCTACCCAACCGCACCCCTACTCACTA
CTGCACAGGTCG
CCAAAGACATAGTAGCACTGCTCAGAAAAGGTGATCTTGTTCACGGTG
TGCCTCAGGAAA
CCGTGCTTCAGCATACTGCTGGCATACTTTCTTGCCTCCCTTCGCTCCTT
GAAGCCCTCC
ACGTGTGTGTACAGCCAGTCCACCACATCCGCCCCTGGCCACAGGTCC
ATCAAAGTCAGG
GTAGCTGAGCCCTGGGAAGCTACGCCAGAATGAGGAACAGACGGGGC
CCTTCCCACACAG

CCAGGGACTCACCAATGACAGCATTGGCAATGGTGATCTTAAGCCACA
TGCGGTCCCGGA
TCTCCAGTCCTGAGTCTGGCAACTGCATGACGCGGACAATGGCACTCA
TGTCACTCTTCA
CAGTCAGCGGTGCCTCCTCAAGCTCTGCAGAGCACACTTCCCTGAGCC
CAGGCTCACAGC
GTGAACCTCCATGGGGTTGAGAGCAGGGGCCAGGGTCAAACCTCTTAT
CTCCCATCCTTG
GGAGATGCCCCTCATCGAAACTTGAGCTAAGACCGGGAGATTCTTCCC
CGTCCCACAGTG
CAAGTCCACGTAGGCAAGGCAGCCCCCCTCCCCTCCCCGGAGAGAACA
AGCTGTTAGCTA
TGTTAGGTAGCAGAAAAGCAAAGCAGAGGCTGCCATGTCCTCCCAATT
CCCCCCTCCGCA
CAGGCCTGGCAGGACCCTCAATTCATGCAGATGACCAGTATGGCCAGG
CCTGGAGGGATA
TGTACATGTATCTTTGTGTACACATTTGTGAAGGTGTTGGAAGCAAAC
AAAACCTTCATA
TGTAATGGGCCCCTGTAATAGCTCTGATGAGCACCAAAGCTCAAAGCT
AGAACTGACCAT
TGTCCTTCAACCTCAGTTTCCTTGGGTGGGGGGGGGTCCTGTGAGCTGC
CACTTACGTGG
GGCGCCAGGCACTGAGCTGGTTAGTGAGGAAGAGCTGGTGCGTGTGAT
GGCGCTGGAGCA
GGGACTCGTACCATAGCGGGGCAGGGCACCCGTCAGTGCTGCTGTGTG
GGACAGCCAGGC
AGCCGGGTCGATGGGTCGCACTGGGTCAGCTGCATAGTTTCCACAGCA
ACGGATTACAGG
TGGTAAGTAGGGGGGCAGCACAGAGGCAGACAAGAAAGACCCCCAGA
CTGAACACAGAAA
CCCCACCCTACCCCACCTTTCCATGGGGTAACTCACCCCTTGGGATGGT
GAAGTAGCTCC

GAGGGGTTGGGTCCCAGCACTTGGCCACTGTGAGACTGATGGGCCTAC
AGAGTTGAGCAG
ACCATGTTGTAAGTGAGGCCCGCACAGCCCCTCCCATCCTGTGCCACT
CCCACCCCCACT
TGGCTCCCACCTCACCCTGTCTGGGACACGATCTCCCGAAGCACCCGT
ACAGCGTCGTCA
TTGCTCATGTTCTCAAAGTTGACATCGTTCACCTACGGGGTTTGTGGGG
TCAGGGGTTGG
TGGTGGGATGTGGGTGCCTCTTGTCCCCACAGTCCCCACATGGCTCCCA
CCTGCAGCAAC
ATGTCGCCCGGCTCAATGCGGCCATCAGCAGCCACGGCCCCGCCCTTC
ATGATGGATCCA
ATGTAGATGCCGCCATCACCCCGGTCGTTGCTCTGGCCCACGATGCTG
ATGCCCAGGAAG
TGGTGCCTCTCTGCAGGAGGGGCCGTGAGCAGGCCCCCAAAGCTCCCG
AGGCTGTACCCA
CCCCCAGCAGGCACCCACAGCCCACAAGGCCTCACCCATGTTGAGAGT
GACGGTGATGAT
GTTCAGGGACATGGTGGAGTCTGTGATGCTGCTGAAGGAGGATGCCTG
CGGAGGGACCCA
GTGAGGGGCTGTGTGGGCACCATTCAGAGCAGACACCCCACCCACCTG
CTGCCTACCCGG
TCTGTCTGCCTCAAGCGCTGCTTCCGACGACGGCATTTGTGCTTCCGAA
CTAGCCGAGAG
GAGGTGCTCTGCTCTGTGGAGCTGCTCAGCCTGAGGCAGGAGTCAGAA
AAGCACAAACAT
GTATAACCAGCTCGGACGCTCAACTACAAATCTCCAGCACGTACTGAC
ATGTGCACACGT
CACCCACCGGCTCGTATTGTCCTCCTCATCTGAGTCAATAAAGCTGCTA
GATTCAAGCTC
ACTGCTCAGTACAGTGGATGCACTGTCTGGAGGTAGTCCCAGGTCCCG
CCGCCGATCCCC

TCTCGGGTGCCCATTGGTCCGGGCAGCTGTGGGGACAGTAGGGTGGGT
ACGACTGTGGGA
CTTCAGTCCTAACAGAATGCGGGTGGCCTGTGCATTTCAAAGTTTATGC
AGTAACTCTGG
GGCCACAGGGGCTAGGAGTACCAGGCTGGGACCTCTACCCAAGGATC
ACTGCTTGGAAGA
ATATGTGGAATACTTCCAGGCTTGGAGTATACCAAAGGGATACCAAG
GG

The polypeptide sequence of mouse SAC1 (SEQ ID NO: 3) is:

MPALAIMGLSLAAFLELGMGASLCLSQQFKAQGDYILGGLFPLGSTEEAT
LNQRTQPNSIPCNRFSPLGLFLAMAMKMAVEEINNGSALLPGLRLGYDLF
DTCSEPVVTMKSSLMFLAKVGSQSIAAYCNYTQYQPRVLAVIGPHSSELA
LITGKFFSFFLMPQVSYSASMDRLSDRETFPSFFRTVPSDRVQLQAVVTLL
QNFSWNWVAALGSDDDYGREGLSIFSSLANARGICIAHEGLVPQHDTSGQ
QLGKVLDVLRQVNQSKVQVVVLFASARAVYSLFSYSIHHGLSPKVWVAS
ESWLTSDLVMTLPNIARVGTVLGFLQRGALLPEFSHYVETHLALAADPAF
CASLNAELDLEEHVMGQRCPRCDDIMLQNLSSGLLQNLSAGQLHHQIFAT
YAAVYSVAQALHNTLQCNVSHCHVSEHVLPWQLLENMYNMSFHARDLT
LQFDAEGNVDMEYDLKMWVWQSPTPVLHTVGTFNGTLQLQQSKMYWP
GNQVPVSQCSRQCKDGQVRRVKGFHSCCYDCVDCKAGSYRKHPDDFTC
TPCNQDQWSPEKSTACLPRRPKFLAWGEPVVLSLLLLLCLVLGLALAALG
LSVHHWDSPLVQASGGSQFCFGLICLGLFCLSVLLFPGRPSSASCLAQQPM
AHLPLTGCLSTLFLQAAETFVESELPLSWANWLCSYLRGLWAWLVVLLA
TFVEAALCAWYLIAFPPEVVTDWSVLPTEVLEHCHVRSWVSLGLVHITNA
MLAFLCFLGTFLVQSQPGRYNRARGLTFAMLAYFITWVSFVPLLANVQV
AYQPAVQMGAILVCALGILVTFHLPKCYVLLWLPKLNTQEFFLGRNAKK
AADENSGGGEAAQGHNE

The cDNA of human SAC1 (SEQ ID NO: 4) is:

ATGCTGGGCCCTGCTGTCCTGGGCCTCAGCCTCTGGGCTCTCCTGCACC
CTGGGACGGGGGCCCCATTGTGCCTGTCACAGCAACTTAGGATGAAGG

GGGACTACGTGCTGGGGGGGCTGTTCCCCCTGGGCGAGGCCGAGGAG
GCTGGCCTCCGCAGCCGGACACGGCCCAGCAGCCCTGTGTGCACCAGG
TTCTCCTCAAACGGCCTGCTCTGGGCACTGGCCATGAAAATGGCCGTG
GAGGAGATCAACAACAAGTCGGATCTGCTGCCCGGGCTGCGCCTGGGC
TACGACCTCTTTGATACGTGCTCGGAGCCTGTGGTGGCCATGAAGCCC
AGCCTCATGTTCCTGGCCAAGGCAGGCAGCCGCGACATCGCCGCCTAC
TGCAACTACACGCAGTACCAGCCCCGTGTGCTGGCTGTCATCGGGCCC
CACTCGTCAGAGCTCGCCATGGTCACCGGCAAGTTCTTCAGCTTCTTCC
TCATGCCCCAGGTCAGCTACGGTGCTAGCATGGAGCTGCTGAGCGCCC
GGGAGACCTTCCCCTCCTTCTTCCGCACCGTGCCCAGCGACCGTGTGCA
GCTGACGGCCGCCGCGGAGCTGCTGCAGGAGTTCGGCTGGAACTGGGT
GGCCGCCCTGGGCAGCGACGACGAGTACGGCCGGCAGGGCCTGAGCA
TCTTCTCGGCCCTGGCCTCGGCACGCGGCATCTGCATCGCGCACGAGG
GCCTGGTGCCGCTGCCCCGTGCCGATGACTCGCGGCTGGGGAAGGTGC
AGGACGTCCTGCACCAGGTGAACCAGAGCAGCGTGCAGGTGGTGCTG
CTGTTCGCCTCCGTGCACGCCGCCCACGCCCTCTTCAACTACAGCATCA
GCAGCAGGCTCTCGCCCAAGGTGTGGGTGGCCAGCGAGGCCTGGCTGA
CCTCTGACCTGGTCATGGGGCTGCCCGGCATGGCCCAGATGGGCACGG
TGCTTGGCTTCCTCCAGAGGGGTGCCCAGCTGCACGAGTTCCCCCAGT
ACGTGAAGACGCACCTGGCCCTGGCCACCGACCCGGCCTTCTGCTCTG
CCCTGGGCGAGAGGGAGCAGGGTCTGGAGGAGGACGTGGTGGGCCAG
CGCTGCCCGCAGTGTGACTGCATCACGCTGCAGAACGTGAGCGCAGGG
CTAAATCACCACCAGACGTTCTCTGTCTACGCAGCTGTGTATAGCGTG
GCCCAGGCCCTGCACAACACTCTTCAGTGCAACGCCTCAGGCTGCCCC
GCGCAGGACCCCGTGAAGCCCTGGCAGCTCCTGGAGAACATGTACAAC
CTGACCTTCCACGTGGGCGGGCTGCCGCTGCGGTTCGACAGCAGCGGA
AACGTGGACATGGAGTACGACCTGAAGCTGTGGGTGTGGCAGGGCTC
AGTGCCCAGGCTCCACGACGTGGGCAGGTTCAACGGCAGCCTCAGGAC
AGAGCGCCTGAAGATCCGCTGGCACACGTCTGACAACCAGAAGCCCGT
GTCCCGGTGCTCGCGGCAGTGCCAGGAGGGCCAGGTGCGCCGGGTCA
AGGGGTTCCACTCCTGCTGCTACGACTGTGTGGACTGCGAGGCGGGCA
GCTACCGGCAAAACCCAGACGACATCGCCTGCACCTTTTGTGGCCAGG

ATGAGTGGTCCCCGGAGCGAAGCACACGCTGCTTCCGCCGCAGGTCTC
GGTTCCTGGCATGGGGCGAGCCGGCTGTGCTGCTGCTGCTCCTGCTGCT
GAGCCTGGCGCTGGGCCTTGTGCTGGCTGCTTTGGGGCTGTTCGTTCAC
CATCGGGACAGCCCACTGGTTCAGGCCTCGGGGGGGCCCCTGGCCTGC
TTTGGCCTGGTGTGCCTGGGCCTGGTCTGCCTCAGCGTCCTCCTGTTCC
CTGGCCAGCCCAGCCCTGCCCGATGCCTGGCCCAGCAGCCCTTGTCCC
ACCTCCCGCTCACGGGCTGCCTGAGCACACTCTTCCTGCAGGCGGCCG
AGATCTTCGTGGAGTCAGAACTGCCTCTGAGCTGGGCAGACCGGCTGA
GTGGCTGCCTGCGGGGGCCCTGGGCCTGGCTGGTGGTGCTGCTGGCCA
TGCTGGTGGAGGTCGCACTGTGCACCTGGTACCTGGTGGCCTTCCCGC
CGGAGGTGGTGACGGACTGGCACATGCTGCCCACGGAGGCGCTGGTG
CACTGCCGCACACGCTCCTGGGTCAGCTTCGGCCTAGCGCACGCCACC
AATGCCACGCTGGCCTTTCTCTGCTTCCTGGGCACTTTCCTGGTGCGGA
GCCAGCCGGGCCGCTACAACCGTGCCCGTGGCCTCACCTTTGCCATGC
TGGCCTACTTCATCACCTGGGTCTCCTTTGTGCCCCTCCTGGCCAATGT
GCAGGTGGTCCTCAGGCCCGCCGTGCAGATGGGCGCCCTCCTGCTCTG
TGTCCTGGGCATCCTGGCTGCCTTCCACCTGCCCAGGTGTTACCTGCTC
ATGCGGCAGCCAGGGCTCAACACCCCCGAGTTCTTCCTGGGAGGGGGC
CCTGGGGATGCCCAAGGCCAGAATGACGGGAACACAGGAAATCAGGG
GAAACATGAGTGA

The polypeptide sequence of human SAC1 substantially from the translated region of the human cDNA (SEQ ID NO: 5) is:

MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEA
GLRSRTRPSSPVCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDL
FDTCSEPVVAMKPSLMFLAKAGSRDIAAYCNYTQYQPRVLAVIGPHSSEL
AMVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRVQLTAAAE
LLQEFGWNWVAALGSDDEYGRQGLSIFSALASARGICIAHEGLVPLPRAD
DSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWV
ASEAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALAT
DPAFCSALGEREQGLEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAA
VYSVAQALHNTLQCNASGCPAQDPVKPWQLLENMYNLTFHVGGLPLRF

DSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRTERLKIRWHTSDN
QKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTFC
GQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFV
HHRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHL
PLTGCLSTLFLQAAEIFVESELPLSWADRLSGCLRGPWAWLVVLLAMLVE
VALCTWYLVAFPPEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNATL
AFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVPLLANVQVVLR
PAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQG
QNDGNTGNQGKHE

III. SAC1 Is a G-Protein Coupled Receptor

The evidence that SAC is a G-protein coupled receptor (GPCR) comes from its sequence homology to other GPCR and the structure predicted for the amino acid sequence.

GPCRs (also known as 7-transmembrane receptors) bind extracellular ligands and transduce signals into the cell by coupling to intracellular G-proteins. GPCRs can be subdivided into more than 30 families on the basis of their ligands. Sac is most closely allied by sequence homology with the Ca^{++}-sensing, metabotropic receptors.

Proteins often contain several modules or domains, each with a distinct evolutionary origin and function. When the Sac cDNA sequence is queried against the Conserved Domain Database at NCBI, the following results are obtained:

Sequences producing significant alignments:				Score (bits)	E Value
Gnl	Pfam	pfam01094	ANF_receptor, Receptor family ligand binding region	145	73-36
Gnl	Pfam	pfam00003	7tm_3, 7-transmembrane receptor (metabotropic glutamate family)	87.0	3e-18

Note the ANF_receptor family contains the metabotropic and calcium-sensing families of GCPs.

The closest sequence homology of the mouse SAC gene is to the Ca^{++} sensing receptors, all of which are GCPRs. An alignment between a calcium sensing GPCR (BAA09453) is shown in Fig. 5.

As described above, all GPCRs have a characteristic 7-transmembrane domain. Figure 6 is a plot of the transmembrane domains of SAC1.

Table 1: Genes Predicted From the Sac Nonrecombinant Interval and Expression Data From NCBI

N	Gene or EST	Size (aa)	How Many EST From Tongue?	Suggested Protein Function
1	*Cyclin ania 6a*	~425	0/36	Potentially involved in differentiation and neural plasticity
2	*Slm1*	~189	0/29	Slm-1 is a Src substrate during mitosis
3	AA404005	446	0/61	Expressed in kidney
4	*Disheveled*	769	0/6	Segment polarity gene; knockouts have a behavioral phenotype
5	*Sac*	746[1]	0/0[2]	Sweet receptor
6	Mm.25556	216	0/5	Weakly similar to Physcomitrella patens glyceraldehyde 3-phosphate dehydrogenase in C. elegans
7	Mm.135238	524	0/5	Expressed in mammary gland and spleen
8	AA435261	328	0/1	Expressed in mouse two cell
9	Centaurin beta 2#	791	0/1	Regulators of membrane traffic and the actin cytoskeleton
10	Voltage gated Na^+channel #	170	0/0	Gumarin reduces the perception of sweet, and may work by blocking sodium channels (Fletcher J.I., Chapman B.E., Mackay J.P., Howden M.E., and King G.F. The structure of versutoxin (delta-atracotoxin-Hv1) provides insights into the binding of site 3 neurotoxins to the voltage-gated sodium channel. *Structure*, 1997;5:1525-1535)
11	*Ubc6p*	597	0/32	Essential for the degradation of misfolded and regulated proteins in the endoplasmic reticulum lumen and membrane
12	Mm.29140	402	0/2	Weakly similar to collagen alpha 1(XVIII) chain

The genomic sequence from AF185591 and RPCI-23-118E21, between the markers that flank the *Sac* nonrecombiant interval, was identified. The repetitive and low complexity sequences were removed, using Repeatmasker (Smit F. and Green P. Repeatmasker). The resulting sequence was analyzed by GENSCAN, which predicted 12 proteins. Of these 12 predicted proteins, one GENSCAN predicted protein was a chimera between two genes (cyclin ania 6a and Slm1). (These sequences were separated into their respective sequences.) The predicted proteins were submitted to a TBLASTN search through the *nr* and the *mouse EST* database at NCBI. Of the 12 predicted proteins, four were named genes, two genes were similar to other named genes (Centaurin beta 2 and the voltage gated Na^+ channel) and are denoted with an #. Three of the predicted proteins were represented as ESTs, and had Unigene cluster numbers. The remaining two predicted genes were identical to previously isolated mouse ESTs. When each predicted protein was blasted against the *mouse EST* database, the number of ESTs from tongue were compared with the number from other tissue sources. No ESTs from these genes appeared in the mouse EST database at NCBI.

1 Note that the GENSCAN prediction is not accurate; sequencing of the cDNA indicates *Sac* is 858 aa.

2 Note that TR1-like is expressed in tongue as detected by RT-PCR. Previously named genes are in *italics*, and ESTs or EST clusters in plain text.

IV. The *Sac* Locus and the *Gpr98* Sweet Taste Receptor Gene

A substantial effort has been devoted to positional cloning of a locus on distal Chr 4 with a major effect on sweetener intake. This locus has been previously described as the *Sac* (saccharin preference) locus, and it also explains ~8 % of the phenotypic variance in ethanol preferences within the B6 × 129 F_2 generation.

Details on positional cloning of the *Sac* locus are found above.

The effects of SAC1 (*Gpr98*) on ethanol intake Two lines of evidence point to the involvement of *Gpr98* in ethanol intake. First, 129.B6-*Sac* congenic mice homozygous for a 194-kb donor fragment from the B6 strain consumed more 10% ethanol solution than did congenic mice without the donor fragment (1.50 ± 0.15 and 1.19 ± 0.11 mL/day, respectively; $p < 0.05$, one-tailed *t*-test). Second, ethanol preference was related to sequence variations of *Gpr98*. Analysis of *Gpr98* sequences from genealogically remote or unrelated mouse strains indicated the presence of two haplotypes of single nucleotide polymorphisms within the *Gpr98* locus. One, 'B6-like' haplotype, was found in mouse strains with elevated sweetener preference and the other, '129-like' haplotype, was found in strains relatively indifferent to sweeteners as described above. Preferences for 10% ethanol for the same mouse strains were studied as described in Abstr. of the 23th RSA Meeting (June 2000, Denver, Colorado). We found that strains with the 'B6-like' haplotype had higher preferences for 10% ethanol ($20 \pm 4\%$, n = 14, strains C57BL/6J, C57L/J, CAST, FVB/NJ, KK/HIJ, NOD/LtJ, NZB/B1NJ, P/J, RBF/DnJ, RF/J, SEA/GnJ, SJL/J, SPRET/Ei and SWR/J) compared with strains having the '129-like' haplotype ($12 \pm 2\%$, n = 10, $p < 0.05$, one-tailed *t*-test, strains 129P3/J, AKR/J, BALB/c, BUB/BnJ, C3H/HeJ, CBA/J, DBA/2J, LP/J, PL/J and RIIIS/J).

V. Preparation of Recombinant or Chemically Synthesized Nucleic Acids, Vectors, Transformation, Host-Cells

Large amounts of the polynucleotides of the present invention may be produced by replication in a suitable host cell. Natural or synthetic polynucleotide

fragments coding for a desired fragment will be incorporated into recombinant polynucleotide constructs, usually DNA constructs, capable of introduction into and replication in a prokaryotic or eukaryotic cell. Usually the polynucleotide constructs will be suitable for replication in a unicellular host, such as yeast or bacteria, but may also be intended for introduction to (with and without integration within the genome) cultured mammalian or plant or other eukaryotic cell lines. The purification of nucleic acids produced by the methods of the present invention is described, e.g., in Ausubel et al., *Current Protocols in Molecular Biology,* Vol. 1-2, John Wiley & Sons, 1992 and Sambrook et al., *Molecular Cloning A Laboratory Manual*, 2nd Ed., Vols. 1-3, Cold Springs Harbor Press, 1989.

The polynucleotides of the present invention may also be produced by chemical synthesis, e.g., by the phosphoramidite method or the triester method, and may be performed on commercial, automated oligonucleotide synthesizers. A double-stranded fragment may be obtained from the single-stranded product of chemical synthesis either by synthesizing the complementary strand and annealing the strands together under appropriate conditions or by adding the complementary strand using DNA polymerase with an appropriate primer sequence.

Polynucleotide constructs prepared for introduction into a prokaryotic or eukaryotic host may comprise a replication system recognized by the host, including the intended polynucleotide fragment encoding the desired polypeptide, and will preferably also include transcription and translational initiation regulatory sequences operably linked to the polypeptide encoding segment. Expression vectors may include, for example, an origin of replication or autonomously replicating sequence (ARS) and expression control sequences, a promoter, an enhancer and necessary processing information sites, such as ribosome-binding sites, RNA splice sites, polyadenylation sites, transcriptional terminator sequences, and mRNA stabilizing sequences. Secretion signals may also be included where appropriate, whether from a native SAC1 protein or from other receptors or from secreted polypeptides of the same or related species, which allow the protein to cross and/or lodge in cell membranes, and thus attain its functional topology, or be secreted from the cell. Such vectors may be prepared by

means of standard recombinant techniques well-known in the art and discussed, for example, in Sambrook et al., 1989 or Ausubel et al., 1992.

An appropriate promoter and other necessary vector sequences will be selected so as to be functional in the host, and may include, when appropriate, those naturally associated with SAC1 genes. Examples of workable combinations of cell lines and expression vectors are described in Sambrook et al., 1989 or Ausubel et al., 1992. Many useful vectors are known in the art and may be obtained from commercial vendors. Promoters such as the trp, lac and phage promoters, TRNA promoters and glycolytic enzyme promoters may be used in prokaryotic hosts. Useful yeast promoters include promoter regions for metallothionein, 3-phosphoglycerate kinase or other glycolytic enzymes such as enolase or glyceraldehyde-3-phosphate dehydrogenase, enzymes responsible for maltose and galactose utilization, and others. In addition, the construct may be joined to an amplifiable gene so that multiple copies of the gene may be made. For appropriate enhancer and other expression control sequences, see also *Enhancers and Eukaryotic Gene Expression*, New York: Cold Spring Harbor Press, 1983. See also, e.g., US Patent Nos. 5,691,198; 5,735,500; 5,747,469 and 5,436,146.

Expression and cloning vectors will likely contain a selectable marker, a gene encoding a protein necessary for survival or growth of a host cell transformed with the vector. The presence of this gene ensures growth of only those host cells which express the inserts. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxic substances, e.g., ampicillin, neomycin, methotrexate, etc.; (b) complement auxotrophic deficiencies; or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli.* The choice of the proper selectable marker will depend on the host cell, and appropriate markers for different hosts are well-known in the art.

The vectors containing the nucleic acids of interest can be transcribed in vitro, and the resulting RNA introduced into the host cell by well-known methods, e.g., by injection, or the vectors can be introduced directly into host cells by methods well-known in the art, which vary depending on the type of cellular host, including electroporation; transfection employing calcium chloride,

rubidium chloride, calcium phosphate, DEAE-dextran, or other substances; microprojectile bombardment; lipofection; infection (where the vector is an infectious agent, such as a retroviral genome); and other methods. The introduction of the polynucleotides into the host cell by any method known in the art, including, inter alia, those described above, will be referred to herein as "transformation." The cells into which have been introduced nucleic acids described above are meant to also include the progeny of such cells.

Large quantities of the nucleic acids and polypeptides of the present invention may be prepared by expressing the SAC1 nucleic acids or portions thereof in vectors or other expression vehicles in compatible prokaryotic or eukaryotic host cells. The most commonly used prokaryotic hosts are strains of *Escherichia coli,* although other prokaryotes, such as *Bacillus subtilis or Pseudomonas* may also be used.

Mammalian or other eukaryotic host cells, such as those of yeast, filamentous fungi, plant, insect, or amphibian or avian species, may also be useful for production of the proteins of the present invention. Propagation of mammalian cells in culture is per se well-known. Examples of commonly used mammalian host cell lines are VERO and HeLa cells, Chinese hamster ovary (CHO) cells, and W138, BHK, and COS cell lines. An example of a commonly used insect cell line is SF9. However, it will be appreciated by the skilled practitioner that other cell lines may be appropriate, e.g., to provide higher expression, desirable glycosylation patterns, or other features.

Clones are selected by using markers depending on the mode of the vector construction. The marker may be on the same or a different DNA molecule, preferably the same DNA molecule. In prokaryotic hosts, the transformant may be selected, e.g., by resistance to ampicillin, tetracycline or other antibiotics. Production of a particular product based on temperature sensitivity may also serve as an appropriate marker.

VI. Diagnosis or Screening

Genetic analysis of obesity and diabetes and alcoholism or alcohol consumption is often complicated by the lack of a simple diagnostic mark. For

example, currently there is no single diagnostic marker for the diagnosis of obesity. Sequence variation of the SAC1 locus may indicate a predisposition to diabetes, obesity, and alcoholism and may provide a diagnostic mark.

In order to detect the presence of a SAC1 allele predisposing an individual to obesity, diabetes, or alcoholism, a biological sample may be prepared and analyzed for the presence or absence of susceptibility alleles of SAC1. Results of these tests and interpretive information may be returned to the health care professionals for communication to the tested individual. Such diagnoses may be performed by diagnostic laboratories. In addition, diagnostic kits may be manufactured and available to health care providers or to private individuals for self-diagnosis.

A basic format for sequence or expression analysis is finding sequences in DNA or RNA extracted from affected family members which create abnormal SAC1 gene products or abnormal levels of SAC1 gene product. The diagnostic or screening method may involve amplification or molecular cloning of the relevant SAC1 sequences. For example, PCR based amplification may be used. Once amplified, the resulting nucleic acid can be sequenced or used as a substrate for DNA probes. Primers and probes specific for the SAC1 gene sequences may be used to identify SAC1 alleles.

The pairs of single-stranded DNA primers can be annealed to sequences within or surrounding the SAC1 gene in order to prime amplifying DNA synthesis of the SAC1 gene itself. The set of primers may allow synthesis of both intron and exon sequences. Allele-specific primers can also be used. Such primers anneal only to particular SAC1 mutant alleles, and thus will only amplify a product in the presence of the mutant allele as a template.

In order to facilitate subsequent cloning of amplified sequences, primers may have restriction enzyme site sequences appended to their 5′ ends. Thus, all nucleotides of the primers are derived from SAC1 sequences or sequences adjacent to SAC1, except for the few nucleotides necessary to form a restriction enzyme site. Such enzymes and sites are well-known in the art. The primers themselves can be synthesized using techniques which are well-known in the art.

Generally, the primers can be made using oligonucleotide synthesizers which are commercially available.

The biological sample to be analyzed, such as blood, may be treated, if desired, to extract the nucleic acids. The sample nucleic acid may be prepared in various ways to facilitate detection of the target sequence; e.g., denaturation, restriction digestion, electrophoresis or dot blotting. The region of interest of the target nucleic acid is usually at least partially single-stranded to form hybrids with the probe. If the sequence is double-stranded, the sequence will probably need to be denatured. The target nucleic acid may be also be fragmented to reduce or eliminate the formation of secondary structures. The fragmentation may be performed using a number of methods, including enzymatic, chemical, thermal cleavage or degradation. For example, fragmentation may be accomplished by heat/Mg^{2+} treatment, endonuclease (e.g., DNAase 1) treatment, restriction enzyme digestion, shearing (e.g., by ultrasound) or NaOH treatment.

Many genotyping and expression monitoring methods have been described previously. In general, target nucleic acid and probe are incubated under conditions which forms hybridization complex between the probe and the target sequence. The region of the probes which is used to bind to the target sequence can be made completely complementary to the targeted region of the SAC1 locus. Therefore, high stringency conditions may be desirable in order to prevent false positives. However, conditions of high stringency are typically used if the probes are complementary to regions of the chromosome which are unique in the genome. The stringency of hybridization is determined by a number of factors during hybridization and during the washing procedure, including temperature, ionic strength, base composition, probe length, and concentration of formamide. Under certain circumstances, the formation of higher order hybrids, such as triplexes, quadraplexes, etc. may be desired to provide the means of detecting target sequences.

Detection, if any, of the resulting hybrid is usually accomplished by the use of labeled probes. Alternatively, the probe may be unlabeled, but may be detectable by specific binding with a ligand which is labeled, either directly or

indirectly. Suitable labels, and methods for labeling probes and ligands are known in the art, and include, for example, radioactive labels which may be incorporated by known methods (e.g., nick translation, random priming or kinase reaction), biotin, fluorescent groups, chemiluminescent groups (e.g., dioxetanes, particularly triggered dioxetanes), enzymes, antibodies and the like. Variations of this basic scheme are known in the art, and include those variations that facilitate separation of the hybrids to be detected from extraneous materials and/or that amplify the signal from the labeled moiety.

Two-step label amplification methodologies are known in the art. These assays work on the principle that a small ligand (such as digoxigenin, biotin, or the like) is attached to a nucleic acid probe capable of specifically binding SAC1.

In one example, the small ligand attached to the nucleic acid probe is specifically recognized by an antibody-enzyme conjugate. In one embodiment of this example, digoxigenin is attached to the nucleic acid probe. Hybridization is detected by an antibody-alkaline phosphatase conjugate which turns over a chemiluminescent substrate. In a second example, the small ligand is recognized by a second ligand-enzyme conjugate that is capable of specifically complexing to the first ligand. A well-known embodiment of this example is the biotin-avidin type of interactions.

It is also contemplated within the scope of this invention that the nucleic acid probe assays of this invention will employ a cocktail of nucleic acid probes capable of detecting SAC1. Thus, in one example to detect the presence of SAC1 in a biological sample, more than one probe complementary to SAC1 is employed.

Predisposition to diabetes, obesity, or alcoholism can be ascertained by testing any fluid or tissue of a human for sequence variations of the SAC1 gene. For example, a person who has inherited a germline SAC1 mutation would be prone to develop obesity, diabetes, or alcoholism. This can be determined by testing DNA from any tissue of the person's body. Most simply, blood can be drawn and DNA extracted from the cells of the blood. In addition, prenatal diagnosis can be accomplished by testing fetal cells, placental cells or amniotic cells for mutations of the SAC1 gene.

The most definitive test for mutations in a candidate locus is to directly compare genomic SAC1 sequences from obese, diabetic, or alcoholic patients, with those from a control population. Alternatively, one could sequence messenger RNA after amplification, e.g., by PCR, thereby eliminating the necessity of determining the exon structure of the candidate gene.

Sequence variations from diabetic, obese, or alcoholic patients falling outside the coding region of SAC1 can be detected by examining the non-coding regions, such as introns and regulatory sequences near or within the SAC1 gene. An early indication that mutations in noncoding regions are important may come from Northern blot experiments that reveal messenger RNA molecules of abnormal size or abundance in obese or diabetic patients as compared to control individuals.

Alteration of SAC1 mRNA expression can be detected by any techniques known in the art (see above). These include Northern blot analysis, PCR amplification, RNase protection, and gene chip analysis. Diminished mRNA expression indicates an alteration of the wild-type SAC1 gene.

The diabetic, obese, or alcoholic condition can also be detected on the basis of the alteration of wild-type SAC1 polypeptide. For example, the presence of a SAC1 gene variant, which produces a protein having a loss of function, or altered function, may directly correlate to an increased risk of obesity or diabetes. Such variation can be determined by sequence analysis in accordance with conventional techniques. For example, antibodies (polyclonal or monoclonal) may be used to detect differences in, or the absence of, SAC1 polypeptides. Antibodies may immunoprecipitate SAC1 proteins from solution as well as react with SAC1 protein on Western or immunoblots of polyacrylamide gels. Antibodies may also detect SAC1 proteins in paraffin or frozen tissue sections, using immunocytochemical techniques. Immunoassay include, for example, enzyme linked immunosorbent assays (ELISA), radioimmunoassays (RIA), immunoradiometric assays (IRMA) and immunoenzymatic assays (IEMA), sandwich assays, etc.

Functional assays, such as protein binding determinations, can be used. Finding a mutant SAC1 gene product indicates alteration of a wild-type SAC1 gene.

VII. Drug, Sweetener, and Alcohol Preference Screening

This invention is also useful for screening compounds by using the SAC1 polypeptide or binding fragment thereof in any of a variety of drug, sweetener, and alcohol screening techniques.

The SAC1 polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, or borne on a cell surface. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant polynucleotides expressing the polypeptide or fragment, preferably in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, for the formation of complexes between a SAC1 polypeptide or fragment and the agent being tested, or examine the degree to which the formation of a complex between a SAC1 polypeptide or fragment and a known ligand is interfered with by the agent being tested.

Thus, the present invention provides methods of screening for drugs and sweeteners comprising contacting such an agent with a SAC1 polypeptide or fragment thereof and assaying (i) for the presence of a complex between the agent and the SAC1 polypeptide or fragment, or (ii) for the presence of a complex between the SAC1 polypeptide or fragment and a ligand, by methods well-known in the art. In such competitive binding assays the SAC1 polypeptide or fragment is typically labeled. Free SAC1 polypeptide or fragment is separated from that present in a protein:protein complex, and the amount of free (i.e., uncomplexed) label is a measure of the binding of the agent being tested to SAC1 or its interference with SAC1:ligand binding, respectively.

Other suitable techniques for drug, sweetener, and alcohol screening may provide high throughput screening for compounds having suitable binding affinity to the SAC1 polypeptides. For example, large numbers of different small peptide

test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with SAC1 polypeptide and washed. Bound SAC1 polypeptide is then detected by methods well-known in the art.

Purified SAC1 can be coated directly onto plates for use in the aforementioned drug screening techniques. However, non-neutralizing antibodies to the polypeptide can be used to capture antibodies to immobilize the SAC1 polypeptide on the solid phase.

This invention also contemplates the use of competitive drug, sweetener, and alcohol screening assays in which neutralizing antibodies capable of specifically binding the SAC1 polypeptide compete with a test compound for binding to the SAC1 polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants of the SAC1 polypeptide.

A further technique for drug, sweetener, and alcohol screening involves the use of host eukaryotic cell lines or cells which have a nonfunctional SAC1 gene. These host cell lines or cells are defective at the SAC1 polypeptide level. The host cell lines or cells are grown in the presence of the drug, sweetener, or alcohol compound. The rate of growth of the host cells is measured to determine if the compound is capable of regulating the growth of SAC1 defective cells.

Briefly, a method of screening for a substance which modulates activity of a polypeptide may include contacting one or more test substances with the polypeptide in a suitable reaction medium, testing the activity of the treated polypeptide and comparing that activity with the activity of the polypeptide in comparable reaction medium untreated with the test substance or substances. A difference in activity between the treated and untreated polypeptides is indicative of a modulating effect of the relevant test substance or substances.

Prior to or as well as being screened for modulation of activity, test substances may be screened for ability to interact with the polypeptide, e.g., in a yeast two-hybrid system. This system may be used as a coarse screen prior to testing a substance for actual ability to modulate activity of the polypeptide.

Alternatively, the screen could be used to screen test substances for binding to a SAC1 specific binding partner, or to find mimetics of a SAC1 polypeptide.

VIII. Rational Drug Design

The goal of rational drug design is to produce structural analogs of biologically active polypeptides of interest or of small molecules with which they interact (e.g., agonists, antagonists, inhibitors) in order to fashion drugs which are, for example, more active or stable forms of the polypeptide, or which, e.g., enhance or interfere with the function of a polypeptide in vivo. In one approach, one first determines the three-dimensional structure of a protein of interest (e.g., SAC1 polypeptide) or, for example, of the SAC1-receptor or ligand complex, by x-ray crystallography, by computer modeling or most typically, by a combination of approaches. Less often, useful information regarding the structure of a polypeptide may be gained by modeling based on the structure of homologous proteins. An example of rational drug design is the development of HIV protease inhibitors. In addition, peptides (e.g., SAC1 polypeptide) are analyzed by an alanine scan. In this technique, an amino acid residue is replaced by Ala, and its effect on the peptide's activity is determined. Each of the amino acid residues of the peptide is analyzed in this manner to determine the important regions of the peptide.

It is also possible to isolate a target-specific antibody, selected by a functional assay, and then to solve its crystal structure. In principle, this approach yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced banks of peptides. Selected peptides would then act as the pharmacore.

Thus, one may design drugs which have, e.g., improved SAC1 polypeptide activity or stability or which act as inhibitors, agonists, antagonists, etc. of SAC1 polypeptide activity. By virtue of the availability of cloned SAC1 sequences,

sufficient amounts of the SAC1 polypeptide may be made available to perform such analytical studies as x-ray crystallography. In addition, the knowledge of the SAC1 protein sequence provided herein will guide those employing computer modeling techniques in place of, or in addition to x-ray crystallography.

Following identification of a substance which modulates or affects polypeptide activity, the substance may be investigated further. Furthermore, it may be manufactured and/or used in preparation, i.e., manufacture or formulation, or a composition such as a medicament, pharmaceutical composition or drug. These may be administered to individuals.

Thus, the present invention extends in various aspects not only to a substance identified using a nucleic acid molecule as a modulator of polypeptide activity, in accordance with what is disclosed herein, but also a pharmaceutical composition, medicament, drug or other composition comprising such a substance, a method comprising administration of such a composition comprising such a substance, a method comprising administration of such a composition to a patient, e.g., for treatment of diabetes, obesity or alcohol consumption, use of such a substance in the manufacture of a composition for administration, e.g., for treatment of diabetes or alcohol consumption, and a method of making a pharmaceutical composition comprising admixing such a substance with a pharmaceutically acceptable excipient, vehicle or carrier, and optionally other ingredients.

A substance identified as a modulator of polypeptide function may be peptide or non-peptide in nature. Non-peptide "small molecules" are often preferred for many in vivo pharmaceutical uses. Accordingly, a mimetic or mimic of the substance (particularly if a peptide) may be designed for pharmaceutical use.

The designing of mimetics to a known pharmaceutically active compound is a known approach to the development of pharmaceuticals based on a "lead" compound. This might be desirable where the active compound is difficult or expensive to synthesize or where it is unsuitable for a particular method of administration, e.g., pure peptides are unsuitable active agents for oral compositions as they tend to be quickly degraded by proteases in the alimentary

canal. Mimetic design, synthesis and testing is generally used to avoid randomly screening large numbers of molecules for a target property.

There are several steps commonly taken in the design of a mimetic from a compound having a given target property. First, the particular parts of the compound that are critical and/or important in determining the target property are determined. In the case of a peptide, this can be done by systematically varying the amino acid residues in the peptide, e.g., by substituting each residue in turn. Alanine scans of peptide are commonly used to refine such peptide motifs. These parts or residues constituting the active region of the compound are known as its pharmacophore.

Once the pharmacophore has been found, its structure is modeled according to its physical properties, e.g., stereochemistry, bonding, size and/or charge, using data from a range of sources, e.g., spectroscopic techniques, x-ray diffraction data and NMR. Computational analysis, similarity mapping (which models the charge and/or volume of a pharmacophore, rather than the bonding between atoms) and other techniques can be used in this modeling process.

In a variant of this approach, the three-dimensional structure of the ligand and its binding partner are modeled. This can be especially used where the ligand and/or binding partner change conformation on binding, allowing the model to take account of this in the design of the mimetic.

A template molecule is then selected onto which chemical groups which mimic the pharmacophore can be grafted. The template molecule and the chemical groups grafted onto it can conveniently be selected so that the mimetic is easy to synthesize, is likely to be pharmacologically acceptable, and does not degrade in vivo, while retaining the biological activity of the lead compound. Alternatively, where the mimetic is peptide-based, further stability can be achieved by cyclizing the peptide, increasing its rigidity. The mimetic(s) found by this approach can then be screened to see whether they have the target property, or to what extent they exhibit it. Further optimization or modification can then be carried out to arrive at one or more final mimetics for in vivo or clinical testing.

IX. Gene Therapy

According to the present invention, a method is also provided of supplying wild-type SAC1 function to a cell which carries mutant SAC1 alleles. The wild-type SAC1 gene or a part of the gene may be introduced into the cell in a vector such that the gene remains extrachromosomal. In such a situation, the gene will be expressed by the cell from the extra chromosomal location. More preferred is the situation where the wild-type SAC1 gene or a part thereof is introduced into the mutant cell in such a way that it recombines with the endogenous mutant SAC1 gene present in the cell. Such recombination requires a double recombination event which results in the correction of the SAC1 gene mutation. Vectors for introduction of genes both for recombination and for extrachromosomal maintenance are known in the art, and any suitable vector may be used. Methods for introducing DNA into cells such as electroporation, calcium phosphate coprecipitation and viral transduction are known in the art, and the choice of method is within the competence of skilled practitioners.

As generally discussed above, the SAC1 gene or fragment, where applicable, may be employed in gene therapy methods in order to increase the amount of the expression products of such genes in diabetic or obese cells. Such gene therapy is particularly appropriate, in which the level of SAC1 polypeptide is absent or compared to normal cells. It may also be useful to increase the level of expression of a given SAC1 gene even in those situations in which the mutant gene is expressed at a "normal" level, but the gene product is not fully functional.

Gene therapy would be carried out according to generally accepted methods, for example, as described by *Therapy for Genetic Diseases*, T. Friedman, ed. Oxford University Press, 1991. Cells from a patient would be first analyzed by the diagnostic methods described above, to ascertain the production of SAC1 polypeptide in these cells. A virus or plasmid vector, containing a copy of the SAC1 gene linked to expression control elements and capable of replicating inside the sample cells, is prepared. Suitable vectors are known, such as disclosed in PCT publications WO 93/07282 and United States

Patent Nos. 5,252,479, 5,691,198, 5,747,469, 5,436,146 and 5,753,500. The vector is then injected into the patient.

Gene transfer systems known in the art may be useful in the practice of the gene therapy methods of the present invention. These include viral and nonviral transfer methods. A number of viruses have been used as gene transfer vectors, including papovaviruses, e.g., SV40, adenovirus, vaccinia virus, adeno-associated virus, herpes viruses including HSV and EBV; lentiviruses, Sindbis and Semliki Forest virus, and retroviruses of avian, murine, and human origin. Most human gene therapy protocols have been based on disabled murine retroviruses.

Nonviral gene transfer methods known in the art include chemical techniques such as calcium phosphate coprecipitation; mechanical techniques, for example microinjection; membrane fusion-mediated transfer via liposomes; and direct DNA uptake and receptor-mediated DNA transfer. Viral-mediated gene transfer can be combined with direct in vivo gene transfer using liposome delivery, allowing one to direct the viral vectors to the affected cells and not into the surrounding nondividing cells. Alternatively, the retroviral vector producer cell line can be injected into affected cells. Injection of producer cells would then provide a continuous source of vector particles.

In an approach which combines biological and physical gene transfer methods, plasmid DNA of any size is combined with a polylysine-conjugated antibody specific to the adenovirus hexon protein, and the resulting complex is bound to an adenovirus vector. The trimolecular complex is then used to infect cells. The adenovirus vector permits efficient binding, internalization, and degradation of the endosome before the coupled DNA is damaged. For other techniques for the delivery of adenovirus based vectors see United States Patent Nos. 5,691,198; 5,747,469; 5,436,146 and 5,753,500.

Liposome/DNA complexes have been shown to be capable of mediating direct in vivo gene transfer. While in standard liposome preparations the gene transfer process is nonspecific, localized in vivo uptake and expression may be accomplished following direct *in situ* administration.

Expression vectors in the context of gene therapy are meant to include those constructs containing sequences sufficient to express a polynucleotide that

has been cloned therein. In viral expression vectors, the construct contains viral sequences sufficient to support packaging of the construct. If the polynucleotide encodes SAC1, expression will produce SAC1. If the polynucleotide encodes an antisense polynucleotide or a ribozyme, expression will produce the antisense polynucleotide or ribozyme. Thus in this context, expression does not require that a protein product be synthesized. In addition to the polynucleotide cloned into the expression vector, the vector also contains a promoter functional in eukaryotic cells. The cloned polynucleotide sequence is under control of this promoter. Suitable eukaryotic promoters include those described above. The expression vector may also include sequences, such as selectable markers and other sequences described herein.

Receptor-mediated gene transfer, for example, may be accomplished by the conjugation of DNA (usually in the form of covalently closed supercoiled plasmid) to a protein ligand via polylysine. Ligands are chosen on the basis of the presence of the corresponding ligand receptors on the cell surface of the target cell/tissue type. One appropriate receptor/ligand pair may include the estrogen receptor and its ligand, estrogen (and estrogen analogues). These ligand-DNA conjugates can be injected directly into the blood if desired and are directed to the target tissue where receptor binding and internalization of the DNA-protein complex occurs. To overcome the problem of intracellular destruction of DNA, coinfection with adenovirus can be included to disrupt endosome function.

X. Peptide Therapy

Peptides which have SAC1 activity can be supplied to cells which carry mutant or missing SAC1 alleles. Protein can be produced by expression of the cDNA sequence in bacteria, for example, using known expression vectors. Alternatively, SAC1 polypeptide can be extracted from SAC1-producing mammalian cells. In addition, the techniques of synthetic chemistry can be employed to synthesize SAC1 protein. Any of such techniques can provide the preparation of the present invention which comprises the SAC1 protein.

Preparation is substantially free of other human proteins. This is most readily accomplished by synthesis in a microorganism or in vitro.

Active SAC1 molecules can be introduced into cells by microinjection or by use of liposomes, for example. Alternatively, some active molecules may be taken up by cells, actively or by diffusion. Extra-cellular application of the SAC1 gene product may be sufficient. Molecules with SAC1 activity (for example, peptides, drugs or organic compounds) may also be used to effect such a reversal. Modified polypeptides having substantially similar function are also used for peptide therapy.

XI. Transformed Hosts

Similarly, cells and animals which carry a mutant SAC1 allele can be used as model systems to study and test for substances which have potential as therapeutic agents. These may be isolated from individuals with SAC1 mutations, either somatic or germline. Alternatively, the cell line can be engineered to carry the mutation in the SAC1 allele.

Animals for testing therapeutic agents can be selected after mutagenesis of whole animals or after treatment of germline cells or zygotes. Such treatments include insertion of mutant SAC1 alleles, usually from a second animal species, as well as insertion of disrupted homologous genes. Alternatively, the endogenous SAC1 gene of the animals may be disrupted by insertion or deletion mutation or other genetic alterations using conventional techniques to produce knockout or transplacement animals. A transplacement is similar to a knockout because the endogenous gene is replaced, but in the case of a transplacement the replacement is by another version of the same gene. After test substances have been administered to the animals, the phenotype must be assessed. If the test substance prevents or suppresses the disease, then the test substance is a candidate therapeutic agent for the treatment of disease. These animal models provide an extremely important testing vehicle for potential therapeutic products.

In one embodiment of the invention, transgenic animals are produced which contain a functional transgene encoding a functional SAC1 polypeptide or

variants thereof. Transgenic animals expressing SAC1 transgenes, recombinant cell lines derived from such animals and transgenic embryos may be useful in methods for screening for and identifying agents that induce or repress function of SAC1. Transgenic animals of the present invention also can be used as models for studying indications such as diabetes.

In one embodiment of the invention, a SAC1 transgene is introduced into a non-human host to produce a transgenic animal expressing a human or murine SAC1 gene. The transgenic animal is produced by the integration of the transgene into the genome in a manner that permits the expression of the transgene. Methods for producing transgenic animals are generally described in US Patent No. 4,873,191 and in *Manipulating the Mouse Embryo; A Laboratory Manual*, 2nd edition (eds., Hogan, Beddington, Costantimi and Long, New York: Cold Spring Harbor Laboratory Press, 1994).

It may be desirable to replace the endogenous SAC1 by homologous recombination between the transgene and the endogenous gene; or the endogenous gene may be eliminated by deletion as in the preparation of "knock-out" animals. Typically, a SAC1 gene flanked by genomic sequences is transferred by microinjection into a fertilized egg. The microinjected eggs are implanted into a host female, and the progeny are screened for the expression of the transgene. Transgenic animals may be produced from the fertilized eggs from a number of animals including, but not limited to reptiles, amphibians, birds, mammals, and fish. Within a particularly preferred embodiment, transgenic mice are generated which express a mutant form of the polypeptide.

As noted above, transgenic animals and cell lines derived from such animals may find use in certain testing experiments. In this regard, transgenic animals and cell lines capable of expressing wild-type or mutant SAC1 may be exposed to test substances. These test substances can be screened for the ability to reduce overexpression of wild-type SAC1 or impair the expression or function of mutant SAC1.

XII. Pharmaceutical Compositions and Routes of Administration

The SAC1 polypeptides, antibodies, peptides and nucleic acids of the present invention can be formulated in pharmaceutical compositions, which are prepared according to conventional pharmaceutical compounding techniques. See, for example, *Remington's Pharmaceutic. Sciences*, 18th Ed. (Easton, PA: Mack Publishing Co., 1990). The composition may contain the active agent or pharmaceutically acceptable salts of the active agent. These compositions may comprise, in addition to one of the active substances, a pharmaceutically acceptable excipient, carrier, buffer, stabilizer or other materials well-known in the art. Such materials should be nontoxic and should not interfere with the efficacy of the active ingredient. The carrier may take a wide variety of forms depending on the form of preparation desired for administration, e.g., intravenous, oral, intrathecal, epineural or parenteral.

For oral administration, the compounds can be formulated into solid or liquid preparations such as capsules, pills, tablets, lozenges, melts, powders, suspensions or emulsions. In preparing the compositions in oral dosage form, any of the usual pharmaceutical media may be employed, such as, for example, water, glycols, oils, alcohols, flavoring agents, preservatives, coloring agents, suspending agents, and the like in the case of oral liquid preparations (such as, for example, suspensions, elixirs and solutions); or carriers such as starches, sugars, diluents, granulating agents, lubricants, binders, disintegrating agents and the like in the case of oral solid preparations (such as, for example, powders, capsules and tablets). Because of their ease in administration, tablets and capsules represent the most advantageous oral dosage unit form, in which case solid pharmaceutical carriers are obviously employed. If desired, tablets may be sugar-coated or enteric-coated by standard techniques. The active agent can be encapsulated to make it stable to passage through the gastrointestinal tract while at the same time allowing for passage across the blood brain barrier. See for example, WO 96/11698.

For parenteral administration, the compound may be dissolved in a pharmaceutical carrier and administered as either a solution or a suspension.

Illustrative of suitable carriers are water, saline, dextrose solutions, fructose solutions, ethanol, or oils of animal, vegetative or synthetic origin. The carrier may also contain other ingredients, for example, preservatives, suspending agents, solubilizing agents, buffers and the like. When the compounds are being administered intrathecally, they may also be dissolved in cerebrospinal fluid.

The active agent is preferably administered in a therapeutically effective amount. The actual amount administered, and the rate and time-course of administration, will depend on the nature and severity of the condition being treated. Prescription of treatment, e.g., decisions on dosage, timing, etc., is within the responsibility of general practitioners or specialists, and typically takes account of the disorder to be treated, the condition of the individual patient, the site of delivery, the method of administration and other factors known to practitioners. Examples of techniques and protocols can be found in *Remington's Pharmaceutical Sciences*.

Alternatively, targeting therapies may be used to deliver the active agent more specifically to certain types of cell, by the use of targeting systems such as antibodies or cell specific ligands. Targeting may be desirable for a variety of reasons, e.g., if the agent is unacceptably toxic, or if it would otherwise require too high a dosage, or if it would not otherwise be able to enter the target cells.

Instead of administering these agents directly, they could be produced in the target cell, e.g., in a viral vector such as described above or in a cell based delivery system such as described in United States Patent No. 5,550,050 and PCT publications WO 92/19195, WO 94/25503, WO 95/01203, WO 95/05452, WO 96/02286, WO 96/02646, WO 96/40871, WO 96/40959 and WO 97/12635, designed for implantation in a patient. The vector could be targeted to the specific cells to be treated, or it could contain regulatory elements which are more tissue specific to the target cells. The cell based delivery system is designed to be implanted in a patient's body at the desired target site and contains a coding sequence for the active agent. Alternatively, the agent could be administered in a precursor form for conversion to the active form by an activating agent produced in, or targeted to, the cells to be treated. See for example, EP 425,731A and WO 90/07936.

EXAMPLES

The following examples further illustrate the present invention. These examples are intended merely to be illustrative of the present invention and are not to be construed as being limiting.

EXAMPLE 1

Animal care and maintenance. All animal protocols used in these studies were approved by the Monell Institutional Animal Care and Use Committee. Mice were housed in individual cages in a temperature- controlled room at 23°C on a 12-hour light:12-hour dark cycle. The animals had free access to deionized water and Teklad Rodent Diet 8604 (Harlan Teklad, Madison, WI).

EXAMPLE 2

Breeding of F2 and partially congenic mice. C57BL6/ByJ (B6) and 129P3/J (formerly named 129/J; abbreviated here as 129) mice were purchased from The Jackson Laboratory. The B6 and 129 mice were outcrossed to produce the first filial generation of hybrids (F_1), and these were intercrossed to produce the second hybrid generation (F_2, n = 629).

To create the partially congenic lines, the F_2 mice were genotyped with several markers on the distal part of chromosome 4, and a few F_2 mice with recombinations in this region were used as founders of strains partially congenic with the 129 strain. These F_2 founders were backcrossed to the 129 strain to produce the N_2 generation. Mice from this and subsequent backcross generations were phenotyped using 96-hour two-bottle tests with saccharin solutions, and genotyped using markers on distal chromosome 4 and on other autosomes. Mice with high saccharin intake (with a fragment of distal chromosome 4 from the B6 strain and homozygous for 129 alleles of markers on other chromosomes) were selected for subsequent backcrossing. This marker-assisted selection resulted in a segregating 129.B6-*Sac* partially congenic strain. Three strains were created, with different overlapping fragments containing the SAC1 gene.

EXAMPLE 3

Testing of sweet preference in the F2 and partially congenic mice. Consumption of 120 mM sucrose and 17 mM saccharin (Sigma Chemical Company, St. Louis, MO) was measured in individually caged mice using 96-hour two-bottle tests, with water as the second choice. The positions of the tubes were switched every 24 hours. Fluid intakes were expressed per 30 g of body weight (the approximate weight of an adult mouse) per day, or as a preference score (ratio of average daily solution intake to total fluid intake, in percent).

EXAMPLE 4

Genotyping of F2 mice and linkage analysis. Genomic DNA was purified from mouse tails by NaOH/Tris (Beier, personal communication; Truett G.E. et al., Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT) [In Process Citation]. *Biotechniques,* 2000;29:52, 54), or the phenol/chloroform method. All F2 mice were genotyped with all available polymorphic microsatellite markers (Research Genetics, Huntsville, AL) known to map near the SAC1 region with a protocol modified slightly from that described by Dietrich W. et al., A genetic map of the mouse suitable for typing intraspecific crosses. *Genetics*, 1991;131;423-447. The markers tested are as follows: D4Mit190, D4Mit42, D4Mit254, and D4Mit256. Analysis of this framework map using MAPMAKER/QTL 1.1 (Lander E. et al. MAPMAKER: An interactive complex package for constructing primary linkage maps of experimental and natural populations. *Genomics,* 1987;1:174-181), indicated that *Sac* mapped distal to D4Mit256, and therefore all available STS and EST were tested by SSCP (Orita M., Iwahana H., Kanazawa H., Hayashi K., and Sekiya T. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphins. *Proceedings of the National Academy of Sciences of the USA*, 1989:86) or direct sequencing, for polymorphisms between the B6 and 129 strains. Polymorphisms between strains were found for the following markers: D18346, AA410003 (K00231), V2r2, and D4Erdt296E, and the linkage analysis conducted again using these polymorphic makers.

EXAMPLE 5

Genotyping of partially congenic mice. Three partially congenic strains of mice were genotyped with all available markers, and those markers with two 129 alleles were excluded from the SAC1 nonrecombinant interval.

EXAMPLE 6

Radiation hybrid mapping. To generate additional markers to narrow the *Sac* nonrecombinant interval, several markers were tested using the T31 RH genome map. Primers from several sequences suggested through survey of the public databases were constructed and DNA from the T31 panel. Results were scored using software at the Jackson Laboratory.

EXAMPLE 7

Construction of BAC contig and marker development. To construct a physical map of the SAC1 region, the RPCI-23 BAC library was screened with markers within and near the SAC1 nonrecombinant interval: each marker was tested by whole cell PCR to confirm its presence. Only those markers positive by both hybridization and PCR are shown. Primers for the BAC ends were constructed from sequence obtained through TIGR (www.tigr.org) or by direct sequencing, when necessary. Each positive clone was tested for the presence of each BAC end (if the BAC end contained unique sequence), and the contig oriented using SEGMAP, Version 3.48 (Green E.D. and Green P. Sequence-tagged site (STS) content mapping of human chromosomes: theoretical considerations and early experiences. *PCR Methods Appl.*, 1991;1:77-90). BAC end sequences was amplified in B6 and 129 strains, and analyzed by SSCP or direct sequencing. Those BAC ends polymorphic between 129 and B6 were tested in the recombinant F2 and partially congenic mice, to further narrow the SAC1 nonrecombinant interval.

EXAMPLE 8

Amplification of SAC1 and polymorphism detection. After the SAC1 nonrecombinant interval was narrowed to less than 350 kb, a 246 kb BAC was chosen for sequencing which spanned most of the region (RPCI-23-118E21). Within this BAC, there was a gene with homology to other taste receptors. Using 11.8 kb of sequence and the program GENSCAN (Burge C.B. and Karlin S. Finding the genes in genomic DNA. *Current Opinion Structural Biology*, 1998;8:346-354), a 858 amino acid protein, with 6 exons, was identified. Primers were constructed that amplified this gene, and an additional 2600 nt upstream and 5200 nt downstream were also amplified (primer sequence available upon request). These PCR products were sequenced using genomic DNA from B6 and 129 mouse strains, as well as other strains with either higher (SWR/J, C57L/J, IS, ST/bJ, SEA/GnJ) or lower (DBA/2J, AKR/J, BALB/cByJ) saccharin preference (Lush I.E., The genetics of tasting in mice. VI. Saccharin, acesulfame, dulcin and sucrose. *Genet Res.*, 1989;53:95-99; Lush I. The genetics of bitterness, sweetness, and saltiness in strains of mice. in *Genetics of perception and communication*, Vol. 3, eds. Wysocki C. and Kare M., New York: Marcel Dekker, 1991:227-235; Lush I.E. and Holland G. The genetics of tasting in mice. V. Glycine and cycloheximide. *Genet Res.*, 1988;52:207-12). Sequences were aligned with Sequencer (Gene Codes, Ann Arbor, MI) and the single nucleotide polymorphisms, insertions and deletions identified.

EXAMPLE 9

Preparation of tongue cDNA and expression studies. Total RNA was extracted from anterior mouse tongue from the 129 and B6 strains (TRIZOL Reagent; GIBCOBRL). Total RNA (200 ng) was reverse transcribed using the Life Technologies SuperScript Kit. Following the reverse transcription, the samples were amplified using Advantage cDNA PCR Kit (Clontech, Palo Alto, CA). Primers were constructed to span exon 2 and 3, so that the genomic and cDNA product size would differ (Primer set 3A;

Left-5′TGCATTGGCCAGACTAGAAA3′;

Right-5CGGCTGGGCTATGACCTAT'). The expected product size for primer 3A is 418 bp for cDNA and 497 bp for genomic DNA. Single bands of these sizes were excised from the gel, purified and sequenced, confirming the intron-exon boundary and expression of mRNA of this gene in mouse tongue. Primers were then designed to cover the whole cDNA, and, the sequences obtained and aligned, to confirm intron/exon boundaries.

EXAMPLE 10

Human gene expression. The human ortholog of the SAC1 gene was examined for mRNA expression in human tongue. Total RNA from human taste papillae was obtained through biopsy, a procedure approved by the Committee on Studies Involving Human Beings at the University of Pennsylvania. The RNA was extracted as described above, reverse transcribed, and amplified, with human specific primers. Two bands were obtained of the expected size for genomic and cDNA. Sequencing of these bands confirmed the SAC1 gene is expressed in human taste papillae.

EXAMPLE 11

Tissue Expression of SAC1. Oligonucleotide primers specific for different parts of the SAC1 gene were used to assay different tissues for SAC1 expression as shown in Table 2. Tissue specific cDNA pools were purchased from OriGene Technologies Ltd. Primer pair 3A, amplifies parts of exons 2 and 3, with a small intron to differentiate between PCR product representing genomic DNA versus cDNA. Primer pair 6A amplifies parts of exons 4 and 5. This part of the protein encodes the 7TM domain, and may cross react with other GPCRs expressed in different tissues.

Table 2. Expression pattern of SAC1

Tissue	3A	6A
Brain	-	-
Heart	-	-
Kidney	+	+
Spleen	+	+
Thymus	+	+
Liver	-	+
Stomach	-	+
Sm Intestine	-	+
Muscle	-	+
Lung	-	+
Testis	+	+
Skin	-	-
Adrenal	+	-
Pancreas	+	+
Uterus	-	-
Prostrate	+	+
Embryo-8.5	-	-
9.5	-	-
12.5	-	-
19	+	-/+
Breast-virgin	-	+
Pregnant	-	+
Lactating	+	+
Involuting	-	-

EXAMPLE 12

Primers for the SAC1 Locus (Seq. ID Nos.: 6-651) are:

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
28.MMHAP7B4.seq	CACTAGAGCTGCCACCTTCC	CCCTCAGCACCACTTTTTGT	162	6-7
28.MMHAP7B4.seq	ACAAAAAGTGGTGCTGAGGG	CAGGAGACCCAAAGGATCAA	163	8-9
AA408705	GCTTCAGAAAATCGAGGCAC	GCATGGGCTATGATAGGTGG	232	10-11
AA408705	TGTTGATCCCACAGCG	CAGGAAATGTCCACTTCTGC		12-13
AA409223	TCTATCTTGCATCCAGCC	GTGCTGTGACTGTGCG		14-15
AA589460	CGCAGCATTTATTTGGAG	CCGACCCTTTAGGAGACAC		16-17
Agrin4	TGTGACTTCCTCTTCCCCAC	TGAGCCACTCCAGATGTCAG	156	18-19
Agrin4	GTGTGTCAGCATCACTGCCT	CCAACGTGCAGTCAAGAAAA	290	20-21
Agrin4	CGAGAGACAAAGTGGTGCTG	TTATGAAGGCCCTCACCAAC	263	22-23
Agrin4	CCAGCTCCTAGAATTGCCTG	GCAGTCTCCCGAAACAAGTC	298	24-25
Agrin4	ATAGAGGAATGGGTGCGATG	TACCAGGAGGGGTCAGTCAG	299	26-27
Agrin4	TACAAGCGAGCTGACCAATG	CCAATCAGCTCGAGTTAGCC	271	28-29
AgrinA	TGCCATTGTGGATGTTCACT	GAGTCCGAGGTCGGTCAATA	575	30-31

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
AgrinB	GCTGGCTTCTGTAGGTCAGG	TATGAGGGTCAAGGGTCAGG	577	32-33
AgrinC	CGCTTTGGTGAGAACTAGCC	CATGTGGAGTTGTGGGAGTG	573	34-35
AgrinD	AATGGGCAGAAGACAGATGG	TATCAGGGTCTGTGAAGCCC	507	36-37
AgrinE	ATACAGGACCCTTTACCCCG	CAGTGTTTCTAGGTCCCCCA	587	38-39
Agrin	GCCTCTGTCTGCCATCTCTC	ATAATGTTACCTGCAGGCGG	594	40-41
AI115523	CTGGAAACACCCATGTCCTC	CGGGCACATGGACACTTTTA	200	42-43
AI225779	GAGCATGAAGTGCAAGGTGA	CGTAGGTGGCACAGTTGAGA	266	44-45
AI225779	GCTGTTAGTGAGGTCAGGGC	CGTAGGTGGCACAGTTGAGA	104	46-47
AI225779	GAGCATGAAGTGCAAGGTGA	TCATTTTCCTAGCCTCGGTG	126	48-49
A022703	TCTAAGAAGATGATGCAGACCC	TGTCCTTCAGGGATAGTGCC		50-51
Cdc2l2	GGCTTCAGCCTCAAGTTCTG	AAAACAACCAAGTTGCCCTG	101	52-53
Cdc2l2	GGCACTGAAATGACCTGGAT	AACAATTCAAGCAACCTCGG	265	54-55
Cdc2l2	CTGTTCCTTCCCAGACTCCA	TTCAGTCACGCAAACCTGAG	225	56-57
Cot	GCCCAGGACTTTGTCACTGT	GGTAACCTGCAGCTCCACTC	284	58-59

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
Cot	GGGACATGCTCTT GGTTCAT	GAACAAAGCCG GGTGATTTA	277	60-61
Cot	GCCCTCAGTTCTC CTAGCCT	GGCAGAGAAGA CTGGTGGAG	110	62-63
Cot	CCCAGACTTAGCG TCTCAGG	AGCAGAGACCTT TGGACTCG	277	64-65
Cot	GAAGGCTGAGTGA GTCCCAG	TTGCACGAGGAG AAGGTTTT	276	66-67
Cot	GATGCCAACGAGA CCTGAAT	AGAAGCCAAAA CCCTCACCT	247	68-69
Cot	AAAAAGCCCTGCA AGAACTT	ATTCAGGTCTCG TTGGCATC	107	70-71
D18346	TGTCCGCAGTGTG GAAACTA	ATGTCCAGGGTA GAGAGCCC	165	72-73
D18402	GGAGTTCTCCTAC CCTGGCT	GAGGCTCTGAGC AGTGTCAA	167	74-75
D4Bir1	GCGATGTTGTTG CG	CAGTGTCTTTCC ACATTT		76-77
D4Ertd296e	AGGCATATTGTAT AATAAATTTGTA GT	CCGGATGACTCT ACTTGAC	201	78-79
D4Hrb1	GCTGTTTATGGGG TCGAGAA	AATTTCTGAAGC AGGGGGAT	194	80-81
D4Hrb1	TCCCCCTGCTTCA GAAATTA	AGGGGGATGATT GTGAGTGA	192	82-83
D4mit313	CTTCTTTAATCAA TCTCTGTCTCTGTG	GGGCACATATGA ACCTCCTG	196	84-85
D4mit344	CCAAACTCTTAGC TTCTTCA	ACACAGAAGAC ACTGAAGAAC	187	86-87

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
D4Mit51	CAGTTGTTAGAAG CAGGATCCC	AGGTGCATATAC CTGGGATACTC	123	88-89
D4Mit59	AGAGTTTGGTCTC TTCCCCTG	TATCCAACACAT TTATGTCTGCG	108	90-91
D4Mit59	GCCAGTGTGCTGA AAGACTG	AGGGACCTGGA GACATCCTT	119	92-93
D4Nds16	CTGTAGGCTGCTT TTATCTTTTG	TGCCCCTTCAGC ACATGCCA		94-95
D4smh6b	TGCAGTGTGACAT GTGCATAGAT	GGAAAGCCAGG CTACGCAGAA	118	96-97
D4smh6b	CTGTAGGCTGCTT TTATCTTTTG	TGCCCCTTCAGC ACATGCCA	102	98-99
D4smh6b	TAGTGTGGTTCCT GACTAACCT	CGGTCTACATAG TGAGTGATTC	181	100-101
D4smh6b	AAAAGCATCCTGC ATCCTTCTG	GGGTTATACAGA GAAACCCTGT	83	102-103
D4Xrf215	TTCCAAGCTCACA CATCAGC	GTGCTGCTCTGC ATTGAGTG	124	104-105
D4Xrf243	GACAGTGTGGGAG AATCCGT	CCCAAGGCATAG GTCACAAT	203	106-107
D4Xrf243	ATTGTGACCTATG CCTTGGG	CGAAGGACCGTC ATCTGAGT	105	108-109
D4Xrf472	GGCTTTGATGTGA AAAAGGC	AGCTCCTCATCG CTCATGTT	245	110-111
D4Xrf472	TGGAACATCTCTG TCGGAAG	GGCTCTCATTGC CACCTTTA	193	112-113
D4X497	CCAGAGAACAGG AGACCTGC	GTGCTGGATACA CTGGCAGA	119	114-115

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
D4Xrf497	GCGAGACGAGTG GGTAGTTC	ACACTGAAACCT CGCTTGCT	129	116-117
D4Xrf497	AGCAAGCGAGGTT TCAGTGT	ACGGGGCTTGAT CCTTTTAT	204	118-119
Dshv4	AAGTTCATGGGCC TCACCACCTGTC	TACTAGCTACCC TTCACATACC	100-300	120-121
Dshv5	ACCTAGCCACTGT CTCAGTCT	ACAGAAGCAGC ATTTACACAG	100-300	122-123
Gnb1	TGGGACAGCTTCC TCAAGAT	AATGGGAATTGT GCTCTTGG	213	124-125
Gnb1	GGGCATCTGGCAA AGATTTA	AGATAACCTGTG TGTCCCGC	281	126-127
Gnb1	GATGTCCGAGAAG GGATGTG	TGTCAGCTTTGA GTGCATCC	277	128-129
Gnb1	ACATGCAGGCTGT TTGACCT	TGTCAGCTTTGA GTGCATCC	166	130-131
K00231	GTGCTCTGCAGAC AAACCAA	GAGCCATTTTGA CCCTTAAA	154	132-133
K00231	TTTCAGGGTCAAA ATGGCTC	TCGACAGCAACT GTGCG		134-135
K00954	GGTGAGAGTGGG GAGATGAA	CCCGGGTGAGTT TAAGAACC	237	136-137
k00954	GGTGAGAGTGGG GAGATGAA	AGGTTAGGCCCA ATTTCCTG	296	138-139
k00954	CCAGGGTTGCTGT ACTGAGA	CAGGTTAGGCCC AATTTCCT	237	140-141
K01153	GGTCAGAGTCCTT CCTTCCC	TCCAACTTCACA GGAAACCC	124	142-143

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
K01153	TTTCCTGTGAAGT TGGAGGG	CACCCATATGGC AAACATCA	213	144-145
K01153	GGTCAGAGTCCTT CCTTCCC	TCCAACTTCACA GGAAACCC	125	146-147
K01153	TGATGTTTGCCAT ATGGGTG	GCTTGCTGCTTC CGATATGT	181	148-149
K01599	GGAAAAGGGAGT CGCCATA	GAGCCGCCTAAC TCTCACAC	166	150-151
K01599	AGGGGATAACCTG CATAGG	ACAAAATTGCTC ATTTGCCC	110	152-153
M-05262	CCATCCCCACTAG CCAGATA	GTCCCCTTTGTC ACAGCAAG	169	154-155
M107-H01	TGAGCACAGGATA GCTCCAC	AAAAGAACACC TGTTTGGGG	217	156-157
M111-B04	TAAACCTCGGCTG TGTGAG	CCCTCAGTGACT TCCTGTGA	267	158-159
M134-C06	CAAAACCACATGG TTACCGA	GCCCTATTGCCA AATGACTT	264	160-161
M134-G01	GGCAGAAAGGAA TCAGAAGC	CACATTAGCCAT TGTCCTGG	161	162-163
M136-B01	TCCTTTATGTCCA ACAGCCA	CATGGTCTGTGA TGTGACCA	164	164-165
M156-H05	ATACCCTTGGTGA GAGCAGG	GCTGTCAAATGA GAAAGGCA	139	166-167
M184-B03	TATTTCATGCTGG GACCAAA	AGAGAAAAACA GTGGGGGTG	89	168-169
Mmp23	CGGGTCCTCTCTT CACCATA	CTACATTTCCCT GAGCTGCC	297	170-171

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
Mmp23	GTTGACCATGTCG GTAACCC	CCACCTCACGGA AACTGAAT	111	172-173
Mmp23	GGTGTTTGGCTCA CAAACCT	GATGCACACACA AAAATCCG	197	174-175
Mmp23	ATCACCCACCAGA ACGAAAA	ACCCTCCAGGAG TAGGTGCT	255	176-177
PCEE	GATGAGACAGTGG GCAAGGT	TTGTCAATAGCA CCAAGCCA	154	178-179
PCEE	GCCTTAATAGCCC CCTTGTT	GCACTCAGCATT GCACAGAT	194	180-181
PCEE	GGACGGACAATTC TGGAAAA	CTATCACACCTC CGATGCCT	142	182-183
PCEE	CAAGCTGGTAGAA TCCCCAA	TCTTTGGAGAAG CAGACCGT	209	184-185
Pkcz	TACAGCATATGCA TGCCAGG	ATTCCTCAGGGC ATTACACG	294	186-187
Pkcz	GCAATCTCTTGTG TCCAGGC	ATTCCTCAGGGC ATTACACG	188	188-189
Pkcz	TACAGCATATGCA TGCCAGG	GGCCTGGACACA AGAGATTG	127	190-191
Pkcz	AAGTGGGTGGACA GTGAAGG	CAGCTTCCTCCA TCTTCTGG	201	192-193
Pkcz	AGAGCCTCCAGTA GATGGCA	TCGTGGACAAGC TCCTTCTT	297	194-195
Pkcz	CATCGAGTATGTC AATGGCG	TTGTCCAGTTTT AGGTCCCG	156	196-197
Pkcz	CAGACTGGGTTTT CCGACAT	GTCAAAGTTGTC CAGGCCAT	132	198-199

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
Pkcz	AGGACGGACCCCA AGATG	TGTCTCGCACTT CCTCACAG	130	200-201
Pkcz	CCAGAAGATGGA GGAAGCTG	TCTACTGGAGGC TCTTGGGA	151	202-203
Pkcz	GAAAAACGACCA GATTTACG	GATCTCAGCAGC ATAGAACC	265	204-205
Pkcz	ACACATTAAGCTG ACGGACT	CAAACATAAGG ACACCCAGT	164	206-207
Pkcz	ACTGGGTGTCCTT ATGTTTG	CCTCTCTTTGGG ATCCTTAT	193	208-209
Pkcz	GTCATAAAGAGGA TCGACCA	GCTCTGTCTAGA AGTGCCTG	252	210-211
Pkcz	ACCAAGACCGAA GAGGGG	GGCATTACACGC TAACTTTTCC	223	212-213
R74924	AGTGCCACCAACC TGGTAAG	AAGTGCCTGCAG GGATGC	165	214-215
R74924	TGCTTTGGTGAGC AATGTTT	AGGGACACCCTT ACCAGGTT	103	216-217
R74924	CTGATGCTTTGGT GAGCAAT	GGGACACCCTTA CCAGGTT		218-219
R75150	ACAGGACAAATGC TGGGTTG	GTGGTAAAGAA CGCTTGGCT	217	220-221
R75150	GGTATCTCACTTG GTAGGAACCTC	AAGAACGCTTGG CTGGC		222-223
RER1 (1)	GCCGATCCTGGTG ATGTACT	ACAATGGCTCAA AACCGTTC		224-225
RER1 (2)	GCCTTGGGAATTT ACCACCT	AGTACATCACCA GGATCGGC		226-227

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
RER1	TAAAAGGCCATGC GATAAGC	AGAGCTCTGTGG GGTTCTCA		228-229
RER1	GAAGGGGACAGT GTTGGAGA	TCCATCAAGGAA GGATCCAC		230-231
Tp73	GGTGGGTAATGAT TGGACT	TGACGTGGAGG GAACTGCC	296-301	232-233
Tp73	TGAGATCTGGTGC CCTCTCT	GCCTGATCTAGG CTGGAAAA	222-229	234-235
Txgp1	AGGCAGAAAGCA GACAAGGA	CGACAGCACTTG TGACCACT	138	236-237
Txgp1	CTGCAGATGTAGA CCAGGCA	CTGTGGTGGATT GGACAGTG	269	238-239
Txgp1	TTGCCTAACACTC CCAAACC	TATTAGGAGCAC CACCAGGC	244	240-241
Txgp1	ACCTGTCTTGTGG GTGGAAG	CTGTGGTGGATT GGACAGTG		242-243
U37351	GTGGCTTGGTGCT ATTGACA	GGGGCTATTAAG GCCATTTT	160	244-245
V2R2	CAATTGAGGAATG GCTACCAA	TGGCTTCATGTC CATTGTGT	170	246-247
V2R2	CAGAACCACAAA GGTAAATTGC	TCATGTTTGCTG TCCAGTTTG	163	248-249
TR1-like1(huma n)	GCCACCATGCTGG GCCCTGCTGTCCT GGG	TCACTCATGTTT CCCCTGATTTCC	2520	250-251
T1-ike2(human)	CTGATTTCCTGTG TTCCCGT	CATGCTGGCCTA CTTCATCA	244	252-253

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
T1-like3(human)	GCCTTGCAGGTCA GCTACGGTGCTAG CAT	TCACTCATGTTT CCCCTGATTTCC	2441	254-255
T1-like4(human)	AGGAAGCAGAGA AAGGCCAG	TCAGAACTGCCT CTGAGCTG	274	256-257
T1-like5(human)	TCTTCACGTACTG GGGGAAC	ACTACAGCATCA GCAGCAGG	175	258-259
T1-like6(human)	AAGCTGAAGAACT TCCCGGT	TGGGCTACGACC TCTTTGAT	211	260-261
h-Tr1like a	ATCTTCAGGCGCT CTGTCCT	GTACGACCTGAA GCTGTGGG		262-263
h-Tr1like b	ATCTTCAGGCGCT CTGTCC	GTACGACCTGAA GCTGTGGG		264-265
h-Tr1like c	ATCTTCAGGCGCT CTGTCC	GAGTACGACCTG AAGCTGTGG		266-267
h-Tr1like d	ATCTTCAGGCGCT CTGTCCT	TACGACCTGAAG CTGTGGG		268-269
h-Tr1like e	ATCTTCAGGCGCT CTGTCC	TACGACCTGAAG CTGTGGG		270-271
h-Tr1like	GCTGTCCCGATGG TGAAC	ACCTTTTGTGGC CAGGATG		272-273
h-Tr1like g	GCTGTCCCGATGG TGAAC	CACCTTTTGTGG CCAGGAT		274-275
h-Tr1like h	GCTGTCCCGATGG TGAAC	CCTTTTGTGGCC AGGATG		276-277
h-Tr1like I	CCTGAACCAGTGG GCTGT	ACCTTTTGTGGC CAGGATG		278-279
h-Tr1like j	CCTGAACCAGTGG GCTGT	CACCTTTTGTGG CCAGGAT		280-281

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
h-Tr1like k	TCATGTTTCCCCT GATTTCC	CATGCTGGCCTA CTTCATCA		282-283
h-Tr1like	ATGAGCAGGTAAC ACCTGGG	TCATCACCTGGG TCTCCTTT		284-285
h-Tr1like m	ATGAGCAGGTAAC ACCTGGG	TTCATCACCTGG GTCTCCTT		286-287
mTr1like-1A	TGGGTTGTGTTCT CTGGTTG	CCTTTTTACAGT CTGCCAGGT		288-289
mTr1like-1B	TGGGTTGTGTTCT CTGGTTG	GATCCCCTTTTT ACAGTCTGC		290-291
mTr1like-2A	ACGGGGTTGGTAC TGTGTGT	CACCCATTGTTA GTGCTGGA		292-293
mTr1like-2B	ACGGGGTTGGTAC TGTGTGT	CACACACCCACC CATTGTTA		294-295
mTr1like-3A	TGCATTGGCCAGA CTAGAAA	CGGCTGGGCTAT GACCTAT		296-297
mTr1like-3B	TGCATTGGCCAGA CTAGAAA	CGGCTGGGCTAT GACCTATT		298-299
mTr1like-4A	GTTCTGCAGCATG ATGTCGT	GGCAGTTGTGAC TCTGTTGC		300-301
mTr1like-4B	GTTCTGCAGCATG ATGTCGT	CTGCAGGCAGTT GTGACTCT		302-303
mTr1like-5A	CCATCCTTTTTGCC TGTCTT	TCTGGAGGAACA TGTGATGG		304-305
mTr1like-5B	CACCATCCTTTTT GCCTGTC	GAACATGTGATG GGGCAAC		306-307
mTr1like-6A	CAAAGCAGCAGG AGGAGTG	AAATGTACTGGC CAGGCAAC		308-309

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
mTr1like-6B	AGTGCTAGACCCA GCACCAG	AAATGTACTGGC CAGGCAAC		310-311
mTr1like-7A	GCACTGACCAGTC TGTCACC	GTCCCCAGAGAA AAGCACAG		312-313
mTr1like-7B	CAGTCTGTCACCA CCTCTGG	CAGTGGTCCCCA GAGAAAAG		314-315
mTr1like-8A	TACTATTCGGGGC TTGTTGG	GCAGCACTATGT GCCTGGTA		316-317
mTr1like-8B	TACTATTCGGGGC TTGTTGG	GCCTGGTATTTG ATCGCTTT		318-319
mTr1like-9A	GCTCAGCTAGGGA TGGAGAA	CAGCTCAGGGAC ACAATGAA		320-321
mTr1like-9B	TCCTACAGGCTAG GGCTCAG	CAGCTCAGGGAC ACAATGAA		322-323
mTr1like-10A	GGGACTGATGTGT GGCTTGT	AGGCGTCCCAGG AATAGAAG		324-325
mTr1like-10B	GGACTGATGTGTG GCTTGTTT	AGGCGTCCCAGG AATAGAAG		326-327
mTr1like-11A	TGTTTCTGTTCTGG TGGCTG	ATCTGCAGGCAG GATCAGAC		328-329
mTr1like-11B	CTCAGTGGTGGGT GACAGTG	ATCTGCAGGCAG GATCAGAC		330-331
Mutation1 (mouse)	ACACACAGTACCA ACCCCGT	CCTGTGGTGATC AAGAAGCA	182	332-333
Mutation2 (mouse)	TGCTTCTTGATCA CCACAGG	GCAACAGAGTC ACAACTGCC	131	334-335
Mutations1+2 (mouse)	ACACACAGTACCA ACCCCGT	GCAACAGAGTC ACAACTGCC	293	336-337

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
34m15-T7	GGGTTTATGTGGC AAGCACT	ACTCCATTTGCC TTTTGTGG	118	338-339
34m15-SP6	CGCTACTTCGCTT TTATCCG	ATGATGACGTAC GACGACGA	150	340-341
37D20-T7	GAAAACAATCGG GGAGAAGTC	TGAAATTATCAC ACGCCAGG	109	342-343
37D20-T7(3)*	AGTGAGAGGCCCA GTCTCAA	GATCTGATGCCC TCTTCTGC	247	344-345
37D20-SP6	GCTAGCCTTGAAG CCAACAC	TGAACAGCATGC TTACCCAG	122	346-347
49O2-T7	TCCCTAGAGGCCT GTCTGTC	TCGTCTCGGAGC CTCTTCTA	169	348-349
49O2-SP6	GATAGTCCCTTAG CCAGCCC	GCCATAGCTCCT CACTGCTC	218	350-351
73B10-T7	CAGAGTGGGCTCT GGTCTTC	TTGTGTTCAGAT GCTCCTGC	237	352-353
73B10-SP6	TTATTTCTGTGCTA GCCGCC	ATCAAGTCAACG TCCCCAAG	267	354-355
75M14	ACCTGGCCTGTGC TAATCTC	GCACCAACCCTA AGAAAGCA	233	356-357
85G18	TCAGGCTAACCTC AAACTCACA	AAAGAAAAGAA AAGAAAAAGTC AGACA	113	358-359
118E21-T7	CCCAGAACTCCAT CCTCAAA	CCCAACCTGTGG TCAGCTAT	185	360-361
118E21-SP6	GGGGCAGGTGGGT AATAAGT	CAAAAGCCCAA CTCCTTGAG	271	362-363
130A12-T7*	GCTCAGTGGGTAA GAGCACC	CTACCCTGCCGC TAATCTCA	242	364-365

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
130A12-SP6	CAGTTAGCACCCC ACCCTAA	TCTGCACCTCTG TTCACCTG	114	366-367
138D7-T7	ACCTCTAGGGTTT ACGGGGA	CCTCAGGTAGTG CAAGCTCC	199	368-369
139J18-T7	TCAGTTACCAAGG GTTTCGG	ATAGGTTGTCAC AGGCCAGG	122	370-371
139J18-SP6	TCAGTTACCAAGG GTTTCGG	ATAGGTTGTCAC AGGCCAGG	122	372-373
147a15-T7*	GTGGTTGCTGGGA TTTGAAC	CAAGCAACCAA ACAACCAAA	101	374-375
147A15-SP6	TCCGGAGGACCAT AAATCTG	CACAGTCCCAGT CATTCCCT	249	376-377
151E4-T7	GTCCCAAAAGCTA GCACAGG	TCATGAGCCACC ATGTGATT	240	378-379
151E4-SP6	GACCTTCGGAAGA GCAGTTG	AGTGTGTGTCGC CATATCCA	223	380-381
152O3-T7	CCTACTCTCTCTCC CCGCTT	GGAAAATGTTTG GCCTTGAA	142	382-383
152O3-SP6	CTGGAGTGAAAGG CAGGAAG	AGGCGGCACCAT ATGAATAA	537	384-385
153B21SP6	TGAGAGTGGGAAT TCTGTTCA	GGATGTAATTGG TGGCAAGG	202	386-387
153B21T7	CTGTTGGAGGAGG TGGCCTA	TGCTTGTATGTT TTTCCTCGT	113	388-389
159J19SP6	TGAGAGTGCCCTC CTCTTTG	GAACCCCTGACC CCAGAC	200	390-391
159J19T7	TGAAGTGCAGATT TTTACATGG	GTTTTGGGGTGG AAAAGGAT	213	392-393

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
189M12SP6*	CCGTCGACATTTAGGTGACA	GATACTGGGGTGGTGGGTAA	189	394-395
227G4-SP6	CCGTCGACATTTAGGTGACA	CGTCCCAGCTGTGTAACTGA	219	396-397
227G4-T7*	GGAAGCAAATGCTCCACTAAA	TATCCCTAGCCCCTTGTGTG	243	398-399
236C12-SP6	CCGTCGACATTTAGGTGACA	GGGTCCTGTTGGTAGTGACC	209	400-401
238O5T7	TATAAGCAGCCCCTCATTGG	CAGGCCAGACACTGCTTACA	244	402-403
238O5SP6	CCTTGGGATCTGGTGTGACT	TGGGTTTAGAGTACGGCTGG	251	404-405
24718-T7	ACCCATTTCCTAATCCCCTG	ATCTCTCCAGCCCCTCTCAG	177	406-407
280G12-T7*	GGGCTGGGAATTGAACCTAT	TGAATCCCTTACAGCCTTGC	420	408-409
280G12-SP6	GCCCCATAAAATCCACTCCT	GCTCCGGAAGGCTAGAAGAT	233	410-411
284D21-T7	GGTTTGGGAGTGTTAGGCAA	ACTCAGTTGGCCTCTCCTCA	138	412-413
284D21-SP6	ACAGAAATCCCTCATGCGA	TCAGTGTGGACCAGAAAGTCC	105	414-415
298E4	TCTGCAAGTCAGCTCTTGATAA	ACTCATAAGGGTCAAGCTGTCTG	100	416-417
298e4-T7(3)*	TCTCCCCTTTTACCACTCCC	GCAAGGAGTCAAAAACAGCA	180	418-419
307E5	GCTAGTTGGGGAACAAACCA	ACTGCAAATGTCCAACTCCA	149	420-421

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
338N4-T7	CAGTTACACAGCT GGGACGA	GCAAGAGCCTA GCAATCCAC	245	422-423
338N4-SP6	CAGTTTAGCACCC CACCCTA	TCTGCACCTCTG TTCACCTG	115	424-425
348P19-SP6	GGGTTCCACTTGA TGCTGAT	TGGTCTGTTTCC TGGAGCTT	227	426-427
350D2-T7*	TGTAGGGAATGTT TCTGCACC	ACATGGAACAG GATTCTGGC	295	428-429
350D2-SP6	GCAGGCAAACAG ACAGACAA	ATGGGGGATCCC TTACTGAC	217	430-431
360M12-T7	CGGTCAGGAGTAG TGTGGGT	CAGCAGCTGATA TTGAGGCA	123	432-433
360M12-SP6	AATGATGAAGTGT CAGCCTCAG	CAACAGAACTCA AAGCCTGG	100	434-435
382A8-SP6	AGCAGGCACAGGT CTCTTGT	AAGAACAGGAC AGTGGTGGG	202	436-437
382A8-SP6(2)	CAGCGATTGGCTC TTCTCTT	GGGGCTTCCTTT CTGAGGTA	531	438-439
386N4-T7	AGCTCAGGTCCAG CTTGGTA	ATTTTCCCCTCC TGCTTCTC	107	440-441
386N4-SP6	CCAAGCCTCTGCT GGTTATC	TGAGGGTGGAG AATGGAAAG	109	442-443
387-T7	GCCCCATAAAATC CACTCCT	TTGCCTAACACT CCCAAACC	214	444-445
387-SP6	CAGTTACACAGCT GGGACGA	GCAAGAGCCTA GCAATCCAC	245	446-447
388I1	CAGCACCTTCCTC TGGTCTC	TGTCTCCAGAGG TTCTGCCT	137	448-449

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
399I12-T7	TGGTGGTGTAATA CTATTCCTTTG	TCTTTAATTTTT GGCTTTTTGATA CA	102	450-451
399I12-SP6	CAGCTGTGTGCAT GTTGACC	CATCATGAAGAC TCAGGGCA	106	452-453
415A22SP6	GTCCACACCTGGC TTTTGTT	CAGCACTCAGTG AGGTTCCA	199	454-455
415G24SP6	ATGTAATGGAAGG GCTGCTG	CAGCACTCAGTG AGGTTCCA	113	456-457
417B22-SP6	AAACAGGCATGA AACTCAGGA	GGGTATCATTGT CACCTCCA	116	458-459
436P10-T7	CACAGGCCAAGTT GTTGTTG	CAGGGGACCTTC TGAATGAT	115	460-461
438C18-T7	AGCTCAGGTCCAG CTTGGTA	ACCACAAAATTT TCCCCTCC	115	462-463
438C18-SP6	CGGGACCTAAAAC TGGACAA	TGGGGACAGTTA CCAGGAAG	254	464-465
457N22-T7	CCGGAGGACCATA AATCTGA	CCTCAAAAACAA GCCTGAGC	129	466-467
457N22-SP6	CCTTCAGAAATGT GTTTGGACA	TCCTGAGTTCAA ATCCCAGC	252	468-469
472O18	CTTTCCATTCTCCA CCCTCA	AGGTCCTAGGGA GAGGTCCA	260	470-471
D4Mon1	AGGCCTACCCAAG GACATCT	GCAGTGAGCTGC AGAGTTTG	201	472-473
D4Mon2	AGACACCCTAGGT CCTGCTG	TGATCTTTCCAA ACGCATAAGA	151	474-475
D4Mon3	GCAAGCAACCTGA ACATGAA	GCTTACGATGGT CGTGAGGT	188	476-477

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
D4Mon4	ACATGCCTGCCTA TCTTTGC	GGAACCTGTTTT CCATGGTG	197	478-479
D4Mon5	ACCTTGTTCCTGG TGTGAGC	TAGCTGGGACGT GGTATGGT	200	480-481
D4Mon6	CCATGGGAGACCA GAAGGTA	TGAGTGTCCTCT GCCTGATG	206	482-483
D4Mon7	GCGCTGACATCCT CCTATGT	CCCACTATGGTC CCAGAGAA	187	484-485
D4Mon8	TTGCACGTCTTTG TTTCGAG	AAAGGGGAATA GACCTGAGTAG AA	219	486-487
D4Mon9	CCAAGAGTCAGCC TTGGAGT	GGACAGGTAGCT CACCCAAC	200	488-489
Tr1likeu1cDNA mouse	TGCCAGCTTTGGC TATCAT	TTCATTGTGTCC CTGAGCTG		490-491
Tr1likeu2cDNA mouse	AGCTTTGGCTATC ATGGGTCTCAG	ACCACCGCCACT GTTCTCATCT		492-493
Tr1like_A1 (mouse)	TGTGGGGGAAGA ACATAGAA	TGATGTGTGGCT TGTTTCTCTT	5935	494-495
Tr1like_A2 (mouse)	ATAGGTGGGGAG GGAGCTAA	TGATGTGTGGCT TGTTTCTCTT	5903	496-497
TR1 like-2 (human)	TGTGCCTGTCACA GCAACTT	CATGCTAGCACC GTAGCTGA		498-499
TR1 like-3 (human)	GGAGACCTTCCCC TCCTTCT	GCTGTAGTTGAA GAGGGCGT		500-501
TR1 like-4 (human)	GTGCTTGGCTTCC TCCAG	CAGGTCGTACTC CATGTCCA		502-503
TR1 like-5 (human)	TGGAGTACGACCT GAAGCTG	ACTCATCCTGGC CACAAAAG		504-505

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
TR1 like-6 (human)	GAACAGGAGGAC GCTGAGG	CTTTTGTGGCCA GGATGAGT		506-507
TR1 like-7 (human)	TCACCTCACCTGG TTGTCAG	GTACGACCTGAA GCTGTGGG		508-509
TR1 like-8 (human)	GGCTGAGATCACA GGGTTGGGTCACT C	CCGTGCCTGTTG GAAGTTGCCTCT GCC		510-511
118e21-0	AATTCCCAGCAAC CACTCAC	CAGACACTCCAG AAGAGGGC	585	512-513
118e21-1	TGACTGCTCTTCC GAAGGTT	TTTGTGGAATAG CCAAAGCC	588	514-515
118-21-2	TCTCTCCTCTCTTC TCCCCC	AGCAGGGTGCAT CACCTTAT	551	516-517
118e21-3	TAGGAGTGCCCCA TAGGTTG	TCATTGTACCCA GCCAGTCA	518	518-519
118e21-4	AGGACTGAGCCTG GATGAGA	CTGGGCGTTTTG TTTTGTTT	552	520-521
118e21-5	CTTCCTCCTGCAG CTACCAC	ACCCTGCTACAA CGCAGACT	546	522-523
118e21-6	TCCAACCTTGACA CCCATTT	AGCCAGGGCTAC ACAGAGAA	584	524-525
139J18T7(1)	CTGCTTTTCCTCA GCAACTG	ATTCGCCGTTAG AAGCTAGG		526-527
139J18T7(2)	AACTGTACGTGGC TGCTGGT	ATTCGCCGTTAG AAGCTAGG		528-529
Agrin(CA)n	GCCAGGTGACCCT TATGAAA	GAGAGATGGCA GACAGAGGC	271	530-531
Agrin(TG)n	AGCTCTCTGTCCC TGGTGAA	TGCCAACCACTA GCCTCTCT	157	532-533

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
repeat1	CTGAACCCTCCAC TCTCCTG	AGCCAGGGCTAC ACAGAGAA	205	534-535
repeat2	AGCCAGGGCTACA CAGAGAA	ACCCTGCTACAA CGCAGACT	153	536-537
repeat3	GCAAGTTTCAGGA GCTAGGG	CCCCAGAACCAG AGACCATA	166	538-539
repeat4	CTAGGGGACTCTG CCAAGTG	CAAGACACCCA GTCCCAACT	195	540-541
repeat5	TACTTCCCCTTTCC CGAACT	TCCTTGGTGCTT ACCCTCAC	232	542-543
repeat6	TGTTCCTGAGTTC ACAACGC	ATTCCCAGCAAC TACATGGC	269	544-545
repeat7	ACATGTCCACTGT GGCAAAA	TGTCATGAGTTT GAGGCCAG	246	546-547
repeat8	ATCAGACAGCCCA CAACCTC	TATGTGCCACCA CACCTGTC	206	548-549
repeat9	GCTCAAGGAAGG ACACACCT	TGCTCTTAACAT TTTGAGCCAT	201	550-551
repeat10	GCTCAGCCCCTGA ATCAATA	GGGATCTGCCTG TCTTACCA	111	552-553
repeat11	GGAAGGTAGGGC CTGGTAAT	GCTCCAAGATCT GTGCGATT	277	554-555
repeat12	TTAGCGTTAGGGT GAGGGTG	GGAGACTACGG ACTTGTGGC	150	556-557
repeat13	CAGTTCTTCCCGA AAACCAC	TTTCTGGGAACT GAGATGGC	174	558-559
repeat14	GTTGGGGCTGCTC ATAGAAA	GCTGTGGCTCTC TTGGAGTT	422	560-561

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
repeat15	CTCTGATTTCCCA CATGCCT	AAGAGGGAGCA CTGAGGACA	152	562-563
repeat16	CAGCAGCAAATGA CCTTTCA	GAGGCAGGCAG ATTTCTGAG	147	564-565
repeat17	GTTTCACATGTTG TGGTGGC	GGGACCTTTGGG ATAGCATT	131	566-567
repeat18	TCAGACATCTCTG GCCTCCT	TTCACTAAGTTG CCCAGGCT	160	568-569
repeat19	TGCCTTTTTCTCAC ATTGTCTC	TTAGAAGCAGA GGCAGAGGC	250	570-571
repeat20	GACCTTTGGAAGA GCAGTCG	TGGCAGCTCACA ATGTCTTT	296	572-573
SHANRU1	GGTGTGGTGTAGG GGAAGAA	TTTCAACTGCAA ACACAAACAG	301	574-575
SHANRU2	AGGGCCAAGGAA GGAGAAT	GCAAATATATAG GGTACCGAGCTG	203	576-577
SHANRU3	CAGATTCTCCAGC TGTCAGG	CTGTGTTTCCGC ACCAAGT	229	578-579
SHANRU4	CTGCCCGTCCTTA TCTTCTG	ACGCACGCTCAC TCATACAC	289	580-581
SHANRU5	CAGCAGAGGTGAT GGGTTCT	TTGTCACACAGT GGTTAAATGC	203	582-583
SHANRU6	TAGAACCGTGGCT GAGGACT	CCGTAAGATAT GAAAGAACTTG GA	201	584-585
SHANRU7	TAATCCTGGCTTA GCGCTTG	TAGAAAGCACA GGGGACAGG	240	586-587
SHANRU8	CCTTCCTCGTCTG AGCTGTT	TTGGGACGTGAC CTGAGAAT	232	588-589

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
SHANRU9	TATGTGTCTGGCC GTTGTTC	GATGTGGGTGCA GGTGAAG	206	590-591
SHANRU10	CCCCTTCTGGAGT GTCTGAA	TCTAGGCAGGGC TACCTTTTT	263	592-593
SHANRU11	GCTGAGCAGCCTC TAGCAA	ACCATGGCTTTT CCCAGTAA	241	594-595
SHANRU12	CTGTGCCTTTGGT GATCAGA	TGTGGCACTCTA CGGCATAA	261	596-597
SHANRU13	TGCATCACTATTA AGCCTCAACC	AAGAATTTGCAA AGACTGTGAGA	260	598-599
SHANRU14	AGCCAGCGCTACA CAGAGA	CTGGACCTTTGG AAGAGCAG	199	600-601
SHANRU15	GGTGGCTCAAACC ATCCATA	GAGGGCAATGA GCAAAATGT	203	602-603
SHANRU16	GGTCCTGTCTCTG GTTCAGG	TAACACCCACAT CAGGCAAC	201	604-605
SHANRU17	TTTCATTTCCTGGT GTTCCTTT	AAACACAGGCG GAACGATAG	198	606-607
SHANRU18	CTATCGTTCCGCC TGTGTTT	AAGGAAGAGGA TGGAGAAAGA	397	608-609
SHANRU19	CGGGTCTTAATGG AGCAGAG	TCCTCCCCAGTT ACCTAGCA	222	610-611
SHANRU20	CAGCAGGCAAGAT GACCTC	GTCCCTCACCAG CCATGTTA	205	612-613
SHANRU21	AGCCTGGGCTAAG TTGTGTG	TATGGGCCAATG TTGTTCCT	204	614-615
SHANRU22	ATGGTGGCTCACA ACCATCT	TTGTCCTCTGAT TGCAGCAT	193	616-617

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
SHANRU23	CTTGGGTCATCAG GCTTTGT	AAGCTGCCCTGC TCTCTCTA	301	618-619
SHANRU24	ATGCTCAGCCTGC TTTGTTT	GCTGATAGCCCT GGGTTCTA	198	620-621
SHANRU25	TGTACGCACAAAT TGACTTGC	GAATCCACATTG CAAAGCCTA	222	622-623
SHANRU26	CACAGGCAAATGA AGGGAAG	CCAGACTTCTCC AGCTCTCC	187	624-625
SHANRU27	TCCTCGAGAGGCT CTAGGTTT	TGCCTAGTCAAC CACAGGAG	237	626-627
SHANRU28	CCTGTGGTTGACT AGGCAGAA	GCCTGATAGCCT GGAATACA	406	628-629
SHANRU29	AAAGGGATGTGTG GCGTAAG	CAAAACCCAACC TTCTCAGC	195	630-631
SHANRU30	TGCACTGACCGTG ATAGAGG	CGGTGTAGCTCT GGCTGTCT	200	632-633
SHANRU31	CATCTCACCAACT CGCACTT	TTTCTGGGAACA AAGAGGCTA	418	634-635
SHANRU32	GAACCCAAGTGTT GGGGTAA	TGGAAGCCCATC TGTCTCTT	222	636-637
SHANRU33	AAATGCAAGTGGG TGCTTCT	CCAGAAGAGGG CGTCAGAT	187	638-639
SHANRU34	GGTGTGCACCACC ATATTCA	GGGAATTATCAG CCAAAAAGC	201	640-641
SHANRU35	GCCCAACTGAAAG CTCAACT	GGAAGGGGGAT AACAATTGAA	263	642-643
SHANRU36	TGCTAATTTCAAG CACAGTGAGA	AGCTTGACACCT TGACAGCA	369	644-645

Marker	Forward	Reverse	Size, bp	SEQ. ID NO.
SHANRU37	AACCTGCAGAGAG GAGACCA	CTCCAAGGGGA GGACTCATT	201	646-647
SHANRU38	TTCAATTGAGTTT CTCTCCTCTGA	TGCAGGACCAA GAAGTAGGC	200	648-649
SHANRU39	CGAGATCTGATGC CCTCTTC	TGCTGAGAGCAG AAAAGGAA	200	650-651

Although the foregoing invention has been described in some detail by way of illustrating and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claims.

All publications, patents, and web sites are herein incorporated by reference in their entirety to the same extent as if each individual publication, patent, or web site was specifically and individually indicated to be incorporated by reference in its entirety.

CLAIMS

What is claimed is:

1. An isolated polynucleotide comprising a sequence variation of SEQ ID. NO 1, wherein said variation is associated with sensing carbohydrates, other sweeteners, or ethanol.

2. An isolated polynucleotide comprising a sequence variation of SEQ ID. NO 2, wherein said variation is associated with sensing carbohydrates, other sweeteners, or ethanol.

3. An isolated polynucleotide comprising a sequence variation of SEQ ID. NO 4, wherein said variation is associated with altered sensation of carbohydrates, other sweeteners, or ethanol.

4. The polynucleotide of Claim 1 wherein said variation is a missense mutation.

5. The polynucleotide of Claim 4 wherein said variation is a nonsense mutation.

6. An isolated polypeptide comprising a variant form of SEQ ID. NO: 3, wherein said variant form is associated with altered preference for carbohydrates, other sweeteners, or ethanol.

7. An isolated polypeptide comprising a variant form of SEQ ID. NO 5, wherein said variant form is associated with altered preference for carbohydrates, other sweeteners, or ethanol.

8. An isolated polynucleotide having at least 8 contiguous nucleotides of the polynucleotides of any one of the Claims 1-3 wherein said 8 contiguous nucleotides span said variation position.

9. An isolated polypeptide having at least four contiguous amino acids of the polypeptides of Claims 6 or 7 wherein said four contiguous amino acids span said variation position.

10. An isolated polynucleotide wherein said polynucleotide is substantially identical to the polynucleotide of Claim 8.

11. An isolated polypeptide wherein said polypeptide is substantially identical to the polypeptide of Claim 9.

12. An isolated polynucleotide having a sequence which is complementary to the polynucleotide of Claim 8 or 10.

13. A polynucleotide specific for the SAC1 locus wherein said polynucleotide hybridizes, under stringent conditions, to at least 8 contiguous nucleotides of the polynucleotide of Claim 1, 2, 3, or 4.

14. The polynucleotide according to Claim 13 wherein said polynucleotide is selected from the group consisting of SEQ ID. NOS 6-651 and homologous equivalents thereof.

15. A polynucleotide specific for the SAC1 locus wherein said polynucleotide that hybridizes, under stringent conditions, to at least 8 contiguous nucleotides of the polynucleotide of Claim 3.

16. The polynucleotide of Claim 15 wherein said polynucleotide is selected from the group consisting of SEQ ID. NOS 6-651 and homologous equivalents thereof.

17. A kit for the detection of the polynucleotide of any one of Claims 1-5, 8, and 10 comprising a polynucleotide that hybridizes, under stringent conditions, to at least 12 contiguous nucleotides of the polynucleotide of any one of the Claims 1-5, 8, and 10, and instructions relating to detection.

18. An isolated antibody which is immunoreactive to the polypeptide of Claim 9 or 11.

19. A method for analyzing a biomolecule in a biological sample, wherein said method comprising:

a) altering SAC1 activity in a biological sample; and

b) measuring the activity.

20. A method for analyzing a polynucleotide in a biological sample comprising the steps of:

a) contacting a polynucleotide in a biological sample with a probe wherein said probe hybridizes to the polynucleotides of Claim 8 or 10 to form a hybridization complex; and

b) detecting the hybridization complex.

21. A method for analyzing the expression of SAC1 comprising the steps of

a) contacting a biological sample with a probe wherein said probe comprises the polynucleotide of Claim 8 or 10; and

b) detecting the expression of SAC1 mRNA transcript in said sample.

22. The method of Claim 19 wherein said step of measuring is an enzymatic assay.

23. The method of Claim 20 or 21 wherein said probe is immobilized on a solid support.

24. The method according to any one of the Claims 19-23 wherein said sample is derived from blood.

25. The method according to any one of the Claims 19-23 wherein said sample is derived from tongue.

26. The method according to any one of the Claims 19-23 wherein said sample is derived from pancreas.

27. The method according to any one of the Claims 19-23 wherein said sample is derived from a human.

28. A method for identifying susceptibility to obesity or diabetes which comprises comparing the nucleotide sequence of the suspected SAC1 allele with a wild type nucleotide sequence, wherein said difference between the suspected allele and the wild-type sequence identifies a sequence variation of the SAC1 nucleotide sequence.

29. An expression vector comprising the polynucleotide of Claim 3, 8, or 10.

30. A host cell comprising the expression vector of Claim 29.

31. A method of producing a polypeptide comprising culturing the cells of Claim 30 and recovering the polypeptide from the host cell.

32. An isolated polypeptide produced according to Claim 31.

33. A method for conducting a screening assay to identify a molecule which enhances or decreases the SAC1 activity comprising the steps of

a) contacting a biological sample with a molecule wherein said biological sample contains SAC1 activity; and

b) analyzing the SAC1 activity in said sample.

34. A pharmaceutical composition comprising

a) the polynucleotide of Claim 8 or 10, the polypeptide of Claim 9 or 11, the antibody of Claim 18 or the molecule of Claim 18; and

b) a suitable pharmaceutical carrier.

35. A method for treating or preventing obesity, diabetes, or alcoholism associated with expression of SAC1, wherein said method comprises administering to a subject an effective amount of the pharmaceutical composition of Claim 34.

36. A transgenic animal that carries an altered SAC1 allele.

37. The transgenic animal of Claim 36 is a knock out mouse.

38. The polypeptide of Claim 6 or 7, wherein said polypeptide is 7-transmembrane G protein coupled receptor (7TM GPCR).

SEQUENCE LISTING

<110> Bachmanov, Alexander A
Beauchamp, Gary K.
Chatterjee, Aurobindo
De Jong, Pieter J.
Li, Shanru
Li, Xia
Ohmen, Jeffrey D
Reed, Danielle R.
Ross, David
Tordoff, Michael G.

<120> GENE AND SEQUENCE VARIATION ASSOCIATED WITH SENSING CARBOHYDRATE COMPOUNDS AND OTHER SWEETNERS

<130> Gene & Sequence Variation.....

<140>
<141>

<150> 60/200,794
<151> 2000-04-28

<160> 652

<170> PatentIn Ver. 2.1

<210> 1
<211> 2577
<212> DNA
<213> Mouse

<400> 1
atgccagctt tggctatcat gggtctcagc ctggctgctt tcctggagct tgggatgggg 60
gcctctttgt gtctgtcaca gcaattcaag gcacaagggg actacatact gggcgggcta 120
tttcccctgg gctcaaccga ggaggccact ctcaaccaga gaacacaacc caacagcatc 180
ccgtgcaaca ggttctcacc ccttggtttg ttcctggcca tggctatgaa gatggctgtg 240
gaggagatca acaatggatc tgccttgctc cctgggctgc ggctgggcta tgacctattt 300
gacacatgct ccgagccagt ggtcaccatg aaatccagtc tcatgttcct ggccaaggtg 360
ggcagtcaaa gcattgctgc ctactgcaac tacacacagt accaaccccg tgtgctggct 420
gtcatcggcc cccactcatc agagcttgcc ctcattacag gcaagttctt cagcttcttc 480
ctcatgccac aggtcagcta tagtgccagc atggatcggc taagtgaccg ggaaacgttt 540
ccatccttct tccgcacagt gcccagtgac cgggtgcagc tgcaggcagt tgtgactctg 600
ttgcagaact tcagctggaa ctgggtggcc gccttaggga gtgatgatga ctatggccgg 660
gaaggtctga gcatcttttc tagtctggcc aatgcacgag gtatctgcat cgcacatgag 720
ggcctggtgc cacaacatga cactagtggc caacagttgg gcaaggtgct ggatgtacta 780
cgccaagtga accaaagtaa agtacaagtg gtggtgctgt ttgcctctgc ccgtgctgtc 840

tactcccttt ttagttacag catccatcat ggcctctcac ccaaggtatg ggtggccagt 900
gagtcttggc tgacatctga cctggtcatg acacttccca atattgcccg tgtgggcact 960
gtgcttgggt ttttgcagcg gggtgcccta ctgcctgaat tttcccatta tgtggagact 1020
caccttgccc tggccgctga cccagcattc tgtgcctcac tgaatgcgga gttggatctg 1080
gaggaacatg tgatggggca acgctgtcca cggtgtgacg acatcatgct gcagaaccta 1140
tcatctgggc tgttgcagaa cctatcagct gggcaattgc accaccaaat atttgcaacc 1200
tatgcagctg tgtacagtgt ggctcaagcc cttcacaaca ccctacagtg caatgtctca 1260
cattgccacg tatcagaaca tgttctaccc tggcagctcc tggagaacat gtacaatatg 1320
agtttccatg ctcgagactt gacactacag tttgatgctg aagggaatgt agacatggaa 1380
tatgacctga agatgtgggt gtggcagagc cctacacctg tattacatac tgtgggcacc 1440
ttcaacggca cccttcagct gcagcagtct aaaatgtact ggccaggcaa ccaggtgcca 1500
gtctcccagt gttcccgcca gtgcaaagat ggccaggttc gccgagtaaa gggctttcat 1560
tcctgctgct atgactgcgt ggactgcaag gcgggcagct accggaagca tccagatgac 1620
ttcacctgta ctccatgtaa ccaggaccag tggtccccag agaaaagcac agcctgctta 1680
cctcgcaggc ccaagtttct ggcttggggg gagccagttg tgctgtcact cctcctgctg 1740
ctttgcctgg tgctgggtct agcactggct gctctggggc tctctgtcca ccactgggac 1800
agccctcttg tccaggcctc aggtggctca cagttctgct ttggcctgat ctgcctaggc 1860
ctctttctgcc tcagtgtcct tctgttccca gggcggccaa gctctgccag ctgccttgca 1920
caacaaccaa tggctcacct ccctctcaca ggctgcctga gcacactctt cctgcaagca 1980
gctgagacct ttgtggagtc tgagctgcca ctgagctggg caaactggct atgcagctac 2040
cttcggggac tctgggcctg gctagtggta ctgttggcca cttttgtgga ggcagcacta 2100
tgtgcctggt atttgatcgc tttcccacca gaggtggtga cagactggtc agtgctgccc 2160
acagaggtac tggagcactg ccacgtgcgt tcctgggtca gcctgggctt ggtgcacatc 2220
accaatgcaa tgttagcttt cctctgcttt ctgggcactt tcctggtaca gagccagcct 2280
ggccgctaca accgtgcccg tggtctcacc ttcgccatgc tagcttattt catcacctgg 2340
gtctcttttg tgcccctcct ggccaatgtg caggtggcct accagccagc tgtgcagatg 2400
ggtgctatcc tagtctgtgc cctgggcatc ctggtcacct tccacctgcc caagtgctat 2460
gtgcttcttt ggctgccaaa gctcaacacc caggagttct tcctgggaag gaatgccaag 2520
aaagcagcag atgagaacag tggcggtggt gaggcagctc agggacacaa tgaatga 2577

<210> 2
<211> 11809
<212> DNA
<213> Mouse

<400> 2
atctgagcct tagacacagc actggtgcca ggcaaacact cctgggccta catgcttggg 60
gcctcttcat attccaaaag ctgtctttgg gtaagatgaa gttcctctgg cagtggcatg 120
agtgctgaag gctctttccc tgcccttcac ctgctttctt gatagtctct ctgcatacca 180
aacaggccct tgtctcctgg gaaatggaaa ctatgaaatc aatagctgag gcttctctag 240
gaaagcctgc cctggtcagt acaacctgtt tcacagcttc tatagaatag ttacatcagc 300
cttctgaaga tggcctctta gagcacatgc acccccaaga ttctaagatg tcaatactaa 360
ctgaccaaac catacctctc tagccagccc tgctgctcct gttgtctggt acccaggtga 420
ctgaggacat gactggtgga aggaaactag gccccttttgt ctgtcagatg gccataccca 480
gcatggctga tgcccagtgt ataagaccct acgcttttcc actggtctta atgttaaacc 540
ctaggacagt gtcctcagca tagctggtgt gtgtgaatgc aaactttggg gcatatctct 600
tccattaagc actgtgatat atgtagtatt tccaacaaat aaattatacc tacatgattg 660

ggtatagcat tctgggatgg gtcacaggtg tgtcaggtgc ctaattatgt gggggaagaa 720
catagaaata tataggtggg gagggagcta accctaggaa taaggctaaa gcatgtgtct 780
ccagtcctga agactcaaag ggcaacgtga atcatgagac atgttcagga ctgaaggagt 840
tgccatgtat ctgtccttga tgtatcttaa tcatacatac actatgagat ctgtgttacc 900
tccattttgc aggtgagaaa agaaacacct gaatggccta ccttaaaggg ctaagtggga 960
aaataggtct gaagataacc caggcactgt gtgacaaagc gggaagaaaa ctagagatgc 1020
tttcttcatg gcaacaacct agagggtaca acctagtggt ttcttcttgg tactccactg 1080
tatacacccc atctgcttgg gctgtacatt gtctgaccat gcttataaca aaagtcacat 1140
actactagcc aagactgaga acttagagcg actggccaga aagtaaagat acaacagttg 1200
atatgtgtgc cacacacaga tccatgtgta catgtctatt aattatgtga acgtgctttg 1260
tggacatcct cacaaagcag cagggaaatg caaaggtcat ttccataaca cctgctggac 1320
accatatgac attgagatta ccggggtgcc cattccaaca agagttaata gctcccccta 1380
tgtttgggtg ccagaaacct gatttgttag caatagctcc ctcacatcca gattaagagg 1440
gggatggctt agctagggtt actatgatga aactatgacc aaagcaactt gtgggtaaaa 1500
gggtgtattt ggcttacact tccatatcac ttcatcaaag tgaggacagg aactcaaata 1560
gagtaggaat ttggtgacaa gagctgatgt agaggcaatg cagtggtgcc acttagtggc 1620
gcgctcagtc tgctcccttt cttaatagaa tgcaagacca ccagcccatg ggtggcacca 1680
caatgggacc gggcccttcc ccatcggtca ctaagaaaat gccctacagc cagatcttat 1740
ggagacattt tctcaacgga ggctcactcc tttcagataa ctctatatca aattgacata 1800
aaccagaaca gaggaggagg ctaagaagga aactgccaat tgcatacatg cacacacctg 1860
gccctagcag ctgcaggaag ctatttgttt atggcctttt ctcattttca tggaccagca 1920
tgagcactct gcagagagag atgcctgcat gcctgccaag gcaggagtgc ttacactgaa 1980
ggtcaacagg atggcagggg ggctgcagag cttccaagtg tcagaacccc agcagaagag 2040
ctgagaccct tgcccgagga ctcaggcggg ttgggaaggc caggaaattc agccagagct 2100
cttcttcaga tggggtacca tctgaaggtt agaccagcta gccagctgtt gttgagggac 2160
cacctctgca gcccctacct ttggaagata gaaagtgtct ctgtgacaag tatggccatt 2220
gtgccccctt attccacagt caacagaaac cctggaatcc tgaacacttc tgcagcttct 2280
tttttacagt ctgccaggtt gctctaggaa tgaagggtgc cgagaggctt gggcgtaggc 2340
aggtgacaag accacagtta gtggtcacag ctggcttact ggatcactct tggacagagt 2400
ttgttagata tggagtggag tatacacaag gcatcaggcg ggggatattg aatgtatcac 2460
cggagctcct tggggcttgg cagccaagca cagcagtggt tttgctaaac aaatccacgg 2520
ttccctcccc ttgacgcagt acatctgtgg ctccaacccc acacacccac ccattgttag 2580
tgctggagac ttctacctac catgccagct ttggctatca tgggtctcag cctggctgct 2640
ttcctggagc ttgggatggg ggcctctttg tgtctgtcac agcaattcaa ggcacaaggg 2700
gactacatac tgggcgggct atttcccctg ggctcaaccg aggaggccac tctcaaccag 2760
agaacacaac ccaacagcat cccgtgcaac aggtatggag gctagtagct ggggtgggag 2820
tgaaccgaag cttggcagct ttggctccgt ggtactacca atctgggaag aggtggtgat 2880
cagtttccat gtggcctcag gttctcaccc cttggtttgt tcctggccat ggctatgaag 2940
atggctgtgg aggagatcaa caatggatct gccttgctcc ctgggctgcg gctgggctat 3000
gacctatttg acacatgctc cgagccagtg gtcaccatga aatccagtct catgttcctg 3060
gccaaggtgg gcagtcaaag cattgctgcc tactgcaact acacacagta ccaaccccgt 3120
gtgctggctg tcatcggccc ccactcatca gagcttgccc tcattacagg caagttcttc 3180
agcttcttcc tcatgccaca ggtgagccca cttcctttgt gttctcaacc gattgcaccc 3240
attgagctct catatcagaa agtgcttctt gatcaccaca ggtcagctat agtgccagca 3300
tggatcggct aagtgaccgg gaaacgtttc catccttctt ccgcacagtg cccagtgacc 3360
gggtgcagct gcaggcagtt gtgactctgt tgcagaactt cagctggaac tgggtggccg 3420
ccttagggag tgatgatgac tatggccggg aaggtctgag catcttttct agtctggcca 3480
atgcacgagg tatctgcatc gcacatgagg gcctggtgcc acaacatgac actagtggcc 3540

aacagttggg caaggtgctg gatgtactac gccaagtgaa ccaaagtaaa gtacaagtgg 3600
tggtgctgtt tgcctctgcc cgtgctgtct actccctttt tagttacagc atccatcatg 3660
gcctctcacc caaggtatgg gtggccagtg agtcttggct gacatctgac ctggtcatga 3720
cacttcccaa tattgcccgt gtgggcactg tgcttgggtt tttgcagcgg ggtgccctac 3780
tgcctgaatt ttcccattat gtggagactc accttgccct ggccgctgac ccagcattct 3840
gtgcctcact gaatgcggag ttggatctgg aggaacatgt gatggggcaa cgctgtccac 3900
ggtgtgacga catcatgctg cagaacctat catctgggct gttgcagaac ctatcagctg 3960
ggcaattgca ccaccaaata tttgcaacct atgcagctgt gtacagtgtg gctcaagccc 4020
ttcacaacac cctacagtgc aatgtctcac attgccacgt atcagaacat gttctaccct 4080
ggcaggtaag ggtagggttt tttgctgggt tttgcctgct cctgcaggaa cactgaacca 4140
ggcagagcca aatcttgttg tgactggaga ggccttaccc tgactccact ccacagctcc 4200
tggagaacat gtacaatatg agtttccatg ctcgagactt gacactacag tttgatgctg 4260
aagggaatgt agacatggaa tatgacctga agatgtgggt gtggcagagc cctacacctg 4320
tattacatac tgtgggcacc ttcaacggca cccttcagct gcagcagtct aaaatgtact 4380
ggccaggcaa ccaggtaagg acaagacagg caaaaaggat ggtgggtaga agcttgtcgg 4440
tcttgggcca gtgctagcca aggggaggcc taacccaagg ctccatgtac aggtgccagt 4500
ctcccagtgt tcccgccagt gcaaagatgg ccaggttcgc cgagtaaagg gctttcattc 4560
ctgctgctat gactgcgtgg actgcaaggc gggcagctac cggaagcatc caggtgaacc 4620
gtcttcccta gacagtctgc acagccgggc tagggggcag aagcattcaa gtctggcaag 4680
cgccctcccg cggggctaat gtggagacag ttactgtggg ggctggctgg ggaggtcggt 4740
ctcccatcag cagaccccac attacttttc ttccttccat cactacagat gacttcacct 4800
gtactccatg taaccaggac cagtggtccc cagagaaaag cacagcctgc ttacctcgca 4860
ggcccaagtt tctggcttgg ggggagccag ttgtgctgtc actcctcctg ctgctttgcc 4920
tggtgctggg tctagcactg gctgctctgg ggctctctgt ccaccactgg gacagccctc 4980
ttgtccaggc ctcaggtggc tcacagttct gctttggcct gatctgccta ggcctcttct 5040
gcctcagtgt ccttctgttc ccagggcggc caagctctgc cagctgcctt gcacaacaac 5100
caatggctca cctccctctc acaggctgcc tgagcacact cttcctgcaa gcagctgaga 5160
cctttgtgga gtctgagctg ccactgagct gggcaaactg gctatgcagc taccttcggg 5220
gactctgggc ctggctagtg gtactgttgg ccacttttgt ggaggcagca ctatgtgcct 5280
ggtatttgat cgctttccca ccagaggtgg tgacagactg gtcagtgctg cccacagagg 5340
tactggagca ctgccacgtg cgttcctggg tcagcctggg cttggtgcac atcaccaatg 5400
caatgttagc tttcctctgc tttctgggca ctttcctggt acagagccag cctggccgct 5460
acaaccgtgc ccgtggtctc accttcgcca tgctagctta tttcatcacc tgggtctctt 5520
ttgtgccccc cctggccaat gtgcaggtgg cctaccagcc agctgtgcag atgggtgcta 5580
tcctagtctg tgccctgggc atcctggtca ccttccacct gcccaagtgc tatgtgcttc 5640
tttggctgcc aaagctcaac acccaggagt tcttcctggg aaggaatgcc aagaaagcag 5700
cagatgagaa cagtggcggt ggtgaggcag ctcagggaca caatgaatga ccactgaccc 5760
gtgaccttcc ctttagggaa cctagcccta ccagaaatct cctaagccaa caagccccga 5820
atagtacctc agcctgagac gtgagacact taactataga cttggactcc actgaccttа 5880
gcctcacagt gacccccttcc ccaaaccccc aaggcctgca gtgcacaaga tggaccctat 5940
gagcccacct atcctttcaa agcaagatta tccttgatcc tattatgccc acctaaggcc 6000
tgcccaggtg acccacaaaa ggttctttgg gacttcatag ccatactttg aattcagaaa 6060
ttccccaggc agaccatggg agaccagaag gtactgcttg cctgaacatg cccagccctg 6120
agccctcact cagcaccctg tccaggcgtc ccaggaatag aaggctgggc atgtatgtgt 6180
gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtat gtacgtatgt atgtatgtat 6240
caggacagaa caagaaagac atcaggcaga ggacactcag gaggtaggca acatccagcc 6300
ttctccatcc ctagctgagc cctagcctgt aggagagaac caggtcgccg ccagcacctt 6360
ggacagatca cacacagggt gcgggtcagc accacggcca gcgccagcca cgcgggaccc 6420

ctggaatcag cttctagtac caaggacaga aaagttgccg caaggcccct tactggccag 6480
caccagggac agagccacat gcctaagcgg caagggacaa gagcatcgtc catctgcagg 6540
caggatcaga cccgggtcag ttctggactg gcccccacac ctgaatcccg gagcagctca 6600
gctggagaaa agagaaacaa gccacacatc agtcccataa aattaaacgc ttttttagt 6660
gtttaaaata gcatttacac agaagcagca tttacacaga agcagctcta tgtcaactac 6720
ccagtcactc agactttgac acagtgtcta gtgtagatgt gtggggccgc tgtgccggga 6780
tggcagtggc acatgatgat gggcagccac cagaacagaa acagaacagg gcccagctct 6840
gcagctcttg tgttcactgt cacccaccac tgagactgag acagtggcta ggtgccaggt 6900
ctctctcctg tctctcctac tagctaccct tcacatacct tcagtacaaa ctgtgttgtc 6960
atgtgccaag tagcaggtgg ggaaaggggc atgcaaactg ccccttggg taactagctg 7020
ccacccttag agcaggcagg ctagcaataa ataaataagt tagaccccac ctgggcagcc 7080
agagaggttt gaaggctctg tctaacccct caaaaatccc accttggcct gacaggtgag 7140
gcccatgaac ttagcgacag tcagcctgtg tccctgtgca cagttctgtg aggctttggg 7200
gcaaggggta ccaagagccc aagagagcct ttcttgttct aaatggaggt cacttccaaa 7260
gaagggaacc aggaggtggt ccctgagact tgtgctgagg acttaaagtc agagatgtct 7320
ccttacaaga ctctatagat acttgagctg taccaccatc agcagcccca agagcagaca 7380
aaatgtcaag ccaatatcct ggtggtatgg ctgccctcag gccctcctct gtagcctgct 7440
ccctctgccc tggcccagag cccacagctg atctatcctg gctggccacc accacggcca 7500
gcgcagagct cctggcacag caggagcaca gactcagcca caggcagcgc tgaagacatt 7560
ggttgatcat cacatgatgt ccacaaagaa ctcacagggg tttcccatgg ccttttggaa 7620
ggactggcgg ctacctgtaa gttctggagg gacagcagcc agctcccgga cgggtggccc 7680
tccaggtggc ccacccacta ctgcataggc ctttgtaagg gggtgcagtg gggggagccc 7740
tggggcaaca gctgaagcct gacttcgagg gctactgcca cggctaagct ggctgacagg 7800
ccgctcccac cagccggtgc taccagaccc acttggtact gtgtggtctg attcactgcc 7860
actaccccca gctccagttg cccggcgctc ctctcggcct ggggtccgat ggctgctccg 7920
tgtggaccca ctgctcttgc tccctagggg gagggaaggg gacaacagag tcagcacgag 7980
gcctggccac ttccagggcc accagctgct cccagacagt cagggcagga cctggtaagc 8040
ctggagatgg taggggaatg gcagccatgc agataccagg aacagctgag aggcgagaag 8100
ctaggggcag tggcagacag cagggacaac aggggccagc ctggcacccc acacctaacc 8160
ccaatgcttg aaccaagggt taatgttaca gctgagaaac taaaaaccag cgaaggccct 8220
gtgtgcccag cattcccatt agccatcctg ggttcaccac ccaaagaccc aaccagggtc 8280
cacccaaccc caggaccctg gtcatctaat ttgcttagcc cctgtcctga aagtagtggg 8340
aacctgaaaa cacgtgctgg ctggggacat gctgagaggg acacaggggg acctggctta 8400
ccggcccgag agtccactct gctagtcctt cagtctaagg cttgctcagc acaaagcaag 8460
ggatagcaca agtcacacac cagtccagtg ctcaccaatg gctaatagga cgattttggg 8520
ccaagctgag cctgggtaca tgcaagggcc tgtccatggt caggattcac tcgatagctt 8580
ccccttgggc tttgccaccc tctggcccaa cctctcctga gtctttctct ggaccttgta 8640
gcacaagtgt gccccactct gcctaagacc tccacatcag tccatctcct cctgagggac 8700
acccaccctt caagatcttc aatatccctg ggatatgctt taacactgat atgctttaac 8760
agtgttgctt gatactctta tctggcactc tgttgggatg caggctccat aactgataaa 8820
gcccattctc cccctagctt ggggcctaga gagtgcccct acctgctatc agtggttact 8880
ttcattcttg ccatatcatc tcctggcctc ttgcctctgc cacctagcac accaggctgt 8940
cttcctattc tctaacggct tctacccaca tcagcccctc cctgtcccac acactgactc 9000
ttgagatgga acccaccggg actcaaacac acagcaggag cacagaggga agcgtcgggg 9060
ccaggcagag cgtgggagtg ggagggagtg ggaggagggg tggcacgcct ctcaccttca 9120
ctctgctggc tcccagcact gccgctgccg cagctgaagc cagggtcctg gtaagcaggc 9180
gggaagcagg gcgggggtcc tgggtactgg taggggtagc cttgacccaa gggccagggt 9240
actgatgggt ggggcagtgg ggccagtgtg tcctgatctg aggctccact ggagccactg 9300

ttgaggttca gggatgcgag gtctggcagg gagggaggga gggaggggta agtgaaggca 9360
aatgaatgag gccacagcaa ccctacccaa ccgcaccct actcactact gcacaggtcg 9420
ccaaagacat agtagcactg ctcagaaaag gtgatcttgt tcacggtgtg cctcaggaaa 9480
ccgtgcttca gcatactgct ggcatacttt cttgcctccc ttcgctcctt gaagccctcc 9540
acgtgtgtgt acagccagtc caccacatcc gcccctggcc acaggtccat caaagtcagg 9600
gtagctgagc cctgggaagc tacgccagaa tgaggaacag acggggccct tcccacacag 9660
ccagggactc accaatgaca gcattggcaa tggtgatctt aagccacatg cggtcccgga 9720
tctccagtcc tgagtctggc aactgcatga cgcggacaat ggcactcatg tcactcttca 9780
cagtcagcgg tgcctcctca agctctgcag agcacacttc cctgagccca ggctcacagc 9840
gtgaacctcc atggggttga gagcaggggc cagggtcaaa cctcttatct cccatccttg 9900
ggagatgccc ctcatcgaaa cttgagctaa gaccgggaga ttcttccccg tcccacagtg 9960
caagtccacg taggcaaggc agccccctc ccctccccgg agagaacaag ctgttagcta 10020
tgttaggtag cagaaaagca aagcagaggc tgccatgtcc tcccaattcc cccctccgca 10080
caggcctggc aggaccctca attcatgcag atgaccagta tggccaggcc tggagggata 10140
tgtacatgta tctttgtgta cacatttgtg aaggtgttgg aagcaaacaa aaccttcata 10200
tgtaatgggc ccctgtaata gctctgatga gcaccaaagc tcaaagctag aactgaccat 10260
tgtccttcaa cctcagtttc cttgggtggg gggggtcct gtgagctgcc acttacgtgg 10320
ggcgccaggc actgagctgg ttagtgagga agagctggtg cgtgtgatgg cgctggagca 10380
gggactcgta ccatagcggg gcagggcacc cgtcagtgct gctgtgtggg acagccaggc 10440
agccgggtcg atgggtcgca ctgggtcagc tgcatagttt ccacagcaac ggattacagg 10500
tggtaagtag gggggcagca cagaggcaga caagaaagac ccccagactg aacacagaaa 10560
ccccacccta ccccaccttt ccatggggta actcaccct tgggatggtg aagtagctcc 10620
gaggggttgg gtcccagcac ttggccactg tgagactgat gggcctacag agttgagcag 10680
accatgttgt aagtgaggcc cgcacagccc ctcccatcct gtgccactcc caccccact 10740
tggctcccac ctcaccctgt ctgggacacg atctcccgaa gcacccgtac agcgtcgtca 10800
ttgctcatgt tctcaaagtt gacatcgttc acctacgggg tttgtggggt caggggttgg 10860
tggtgggatg tgggtgcctc ttgtccccac agtccccaca tggctcccac ctgcagcaac 10920
atgtcgcccg gctcaatgcg gccatcagca gccacggccc cgcccttcat gatggatcca 10980
atgtagatgc cgccatcacc ccggtcgttg ctctggccca cgatgctgat gcccaggaag 11040
tggtgcctct ctgcaggagg ggccgtgagc aggcccccaa agctcccgag gctgtaccca 11100
cccccagcag gcacccacag cccacaaggc ctcacccatg ttgagagtga cggtgatgat 11160
gttcagggac atggtggagt ctgtgatgct gctgaaggag gatgcctgcg gagggaccca 11220
gtgaggggct gtgtgggcac cattcagagc agacacccca cccacctgct gcctacccgg 11280
tctgtctgcc tcaagcgctg cttccgacga cggcatttgt gcttccgaac tagccgagag 11340
gaggtgctct gctctgtgga gctgctcagc ctgaggcagg agtcagaaaa gcacaaacat 11400
gtataaccag ctcggacgct caactacaaa tctccagcac gtactgacat gtgcacacgt 11460
cacccaccgg ctcgtattgt cctcctcatc tgagtcaata aagctgctag attcaagctc 11520
actgctcagt acagtggatg cactgtctgg aggtagtccc aggtcccgcc gccgatcccc 11580
tctcgggtgc ccattggtcc gggcagctgt ggggacagta gggtgggtac gactgtggga 11640
cttcagtcct aacagaatgc gggtggcctg tgcatttcaa agtttatgca gtaactctgg 11700
ggccacaggg gctaggagta ccaggctggg acctctaccc aaggatcact gcttggaaga 11760
atatgtggaa tacttccagg cttggagtat accaaaggga taccaaggg 11809

<210> 3
<211> 858
<212> PRT
<213> Mouse

<400> 3

```
Met Pro Ala Leu Ala Ile Met Gly Leu Ser Leu Ala Ala Phe Leu Glu
 1               5                  10                  15

Leu Gly Met Gly Ala Ser Leu Cys Leu Ser Gln Gln Phe Lys Ala Gln
            20                  25                  30

Gly Asp Tyr Ile Leu Gly Gly Leu Phe Pro Leu Gly Ser Thr Glu Glu
        35                  40                  45

Ala Thr Leu Asn Gln Arg Thr Gln Pro Asn Ser Ile Pro Cys Asn Arg
    50                  55                  60

Phe Ser Pro Leu Gly Leu Phe Leu Ala Met Ala Met Lys Met Ala Val
65                  70                  75                  80

Glu Glu Ile Asn Asn Gly Ser Ala Leu Leu Pro Gly Leu Arg Leu Gly
                85                  90                  95

Tyr Asp Leu Phe Asp Thr Cys Ser Glu Pro Val Val Thr Met Lys Ser
            100                 105                 110

Ser Leu Met Phe Leu Ala Lys Val Gly Ser Gln Ser Ile Ala Ala Tyr
        115                 120                 125

Cys Asn Tyr Thr Gln Tyr Gln Pro Arg Val Leu Ala Val Ile Gly Pro
    130                 135                 140

His Ser Ser Glu Leu Ala Leu Ile Thr Gly Lys Phe Phe Ser Phe Phe
145                 150                 155                 160

Leu Met Pro Gln Val Ser Tyr Ser Ala Ser Met Asp Arg Leu Ser Asp
                165                 170                 175

Arg Glu Thr Phe Pro Ser Phe Phe Arg Thr Val Pro Ser Asp Arg Val
            180                 185                 190

Gln Leu Gln Ala Val Val Thr Leu Leu Gln Asn Phe Ser Trp Asn Trp
        195                 200                 205

Val Ala Ala Leu Gly Ser Asp Asp Asp Tyr Gly Arg Glu Gly Leu Ser
    210                 215                 220

Ile Phe Ser Ser Leu Ala Asn Ala Arg Gly Ile Cys Ile Ala His Glu
225                 230                 235                 240

Gly Leu Val Pro Gln His Asp Thr Ser Gly Gln Gln Leu Gly Lys Val
```

245 250 255

Leu Asp Val Leu Arg Gln Val Asn Gln Ser Lys Val Gln Val Val Val
260 265 270

Leu Phe Ala Ser Ala Arg Ala Val Tyr Ser Leu Phe Ser Tyr Ser Ile
275 280 285

His His Gly Leu Ser Pro Lys Val Trp Val Ala Ser Glu Ser Trp Leu
290 295 300

Thr Ser Asp Leu Val Met Thr Leu Pro Asn Ile Ala Arg Val Gly Thr
305 310 315 320

Val Leu Gly Phe Leu Gln Arg Gly Ala Leu Leu Pro Glu Phe Ser His
325 330 335

Tyr Val Glu Thr His Leu Ala Leu Ala Ala Asp Pro Ala Phe Cys Ala
340 345 350

Ser Leu Asn Ala Glu Leu Asp Leu Glu Glu His Val Met Gly Gln Arg
355 360 365

Cys Pro Arg Cys Asp Asp Ile Met Leu Gln Asn Leu Ser Ser Gly Leu
370 375 380

Leu Gln Asn Leu Ser Ala Gly Gln Leu His His Gln Ile Phe Ala Thr
385 390 395 400

Tyr Ala Ala Val Tyr Ser Val Ala Gln Ala Leu His Asn Thr Leu Gln
405 410 415

Cys Asn Val Ser His Cys His Val Ser Glu His Val Leu Pro Trp Gln
420 425 430

Leu Leu Glu Asn Met Tyr Asn Met Ser Phe His Ala Arg Asp Leu Thr
435 440 445

Leu Gln Phe Asp Ala Glu Gly Asn Val Asp Met Glu Tyr Asp Leu Lys
450 455 460

Met Trp Val Trp Gln Ser Pro Thr Pro Val Leu His Thr Val Gly Thr
465 470 475 480

Phe Asn Gly Thr Leu Gln Leu Gln Gln Ser Lys Met Tyr Trp Pro Gly
485 490 495

Asn Gln Val Pro Val Ser Gln Cys Ser Arg Gln Cys Lys Asp Gly Gln

```
          500                 505                 510

Val Arg Arg Val Lys Gly Phe His Ser Cys Cys Tyr Asp Cys Val Asp
        515                 520                 525

Cys Lys Ala Gly Ser Tyr Arg Lys His Pro Asp Asp Phe Thr Cys Thr
    530                 535                 540

Pro Cys Asn Gln Asp Gln Trp Ser Pro Glu Lys Ser Thr Ala Cys Leu
545                 550                 555                 560

Pro Arg Arg Pro Lys Phe Leu Ala Trp Gly Glu Pro Val Val Leu Ser
                565                 570                 575

Leu Leu Leu Leu Leu Cys Leu Val Leu Gly Leu Ala Leu Ala Ala Leu
            580                 585                 590

Gly Leu Ser Val His His Trp Asp Ser Pro Leu Val Gln Ala Ser Gly
        595                 600                 605

Gly Ser Gln Phe Cys Phe Gly Leu Ile Cys Leu Gly Leu Phe Cys Leu
    610                 615                 620

Ser Val Leu Leu Phe Pro Gly Arg Pro Ser Ser Ala Ser Cys Leu Ala
625                 630                 635                 640

Gln Gln Pro Met Ala His Leu Pro Leu Thr Gly Cys Leu Ser Thr Leu
                645                 650                 655

Phe Leu Gln Ala Ala Glu Thr Phe Val Glu Ser Glu Leu Pro Leu Ser
            660                 665                 670

Trp Ala Asn Trp Leu Cys Ser Tyr Leu Arg Gly Leu Trp Ala Trp Leu
        675                 680                 685

Val Val Leu Leu Ala Thr Phe Val Glu Ala Ala Leu Cys Ala Trp Tyr
    690                 695                 700

Leu Ile Ala Phe Pro Pro Glu Val Val Thr Asp Trp Ser Val Leu Pro
705                 710                 715                 720

Thr Glu Val Leu Glu His Cys His Val Arg Ser Trp Val Ser Leu Gly
                725                 730                 735

Leu Val His Ile Thr Asn Ala Met Leu Ala Phe Leu Cys Phe Leu Gly
            740                 745                 750

Thr Phe Leu Val Gln Ser Gln Pro Gly Arg Tyr Asn Arg Ala Arg Gly
```

```
        755                 760                 765

Leu Thr Phe Ala Met Leu Ala Tyr Phe Ile Thr Trp Val Ser Phe Val
    770                 775                 780

Pro Leu Leu Ala Asn Val Gln Val Ala Tyr Gln Pro Ala Val Gln Met
785                 790                 795                 800

Gly Ala Ile Leu Val Cys Ala Leu Gly Ile Leu Val Thr Phe His Leu
                805                 810                 815

Pro Lys Cys Tyr Val Leu Leu Trp Leu Pro Lys Leu Asn Thr Gln Glu
            820                 825                 830

Phe Phe Leu Gly Arg Asn Ala Lys Lys Ala Ala Asp Glu Asn Ser Gly
        835                 840                 845

Gly Gly Glu Ala Ala Gln Gly His Asn Glu
    850                 855
```

<210> 4
<211> 2559
<212> DNA
<213> Homo sapiens

<400> 4

```
atgctgggcc ctgctgtcct gggcctcagc ctctgggctc tcctgcaccc tgggacgggg 60
gccccattgt gcctgtcaca gcaacttagg atgaaggggg actacgtgct gggggggctg 120
ttcccccctgg gcgaggccga ggaggctggc ctccgcagcc ggacacggcc cagcagccct 180
gtgtgcacca ggttctcctc aaacggcctg ctctgggcac tggccatgaa aatggccgtg 240
gaggagatca acaacaagtc ggatctgctg cccgggctgc gcctgggcta cgacctcttt 300
gatacgtgct cggagcctgt ggtggccatg aagcccagcc tcatgttcct ggccaaggca 360
ggcagccgcg acatcgccgc ctactgcaac tacacgcagt accagccccg tgtgctggct 420
gtcatcgggc cccactcgtc agagctcgcc atggtcaccg gcaagttctt cagcttcttc 480
ctcatgcccc aggtcagcta cggtgctagc atggagctgc tgagcgcccg ggagaccttc 540
ccctccttct tccgcaccgt gcccagcgac cgtgtgcagc tgacggccgc cgcggagctg 600
ctgcaggagt tcggctggaa ctgggtggcc gccctgggca gcgacgacga gtacggccgg 660
cagggcctga gcatcttctc ggccctggcc tcggcacgcg gcatctgcat cgcgcacgag 720
ggcctggtgc cgctgccccg tgccgatgac tcgcggctgg ggaaggtgca ggacgtcctg 780
caccaggtga accagagcag cgtgcaggtg gtgctgctgt tcgcctccgt gcacgccgcc 840
cacgccctct tcaactacag catcagcagc aggctctcgc ccaaggtgtg ggtggccagc 900
gaggcctggc tgacctctga cctggtcatg gggctgcccg gcatggccca gatgggcacg 960
gtgcttggct tcctccagag gggtgcccag ctgcacgagt tcccccagta cgtgaagacg 1020
cacctggccc tggccaccga cccggccttc tgctctgccc tgggcgagag ggagcagggt 1080
ctggaggagg acgtggtggg ccagcgctgc ccgcagtgtg actgcatcac gctgcagaac 1140
gtgagcgcag ggctaaatca ccaccagacg ttctctgtct acgcagctgt gtatagcgtg 1200
gcccaggccc tgcacaacac tcttcagtgc aacgcctcag gctgccccgc gcaggacccc 1260
```

gtgaagccct ggcagctcct ggagaacatg tacaacctga ccttccacgt gggcgggctg 1320
ccgctgcggt tcgacagcag cggaaacgtg gacatggagt acgacctgaa gctgtgggtg 1380
tggcagggct cagtgcccag gctccacgac gtgggcaggt tcaacggcag cctcaggaca 1440
gagcgcctga agatccgctg gcacacgtct gacaaccaga agcccgtgtc ccggtgctcg 1500
cggcagtgcc aggagggcca ggtgcgccgg gtcaaggggt tccactcctg ctgctacgac 1560
tgtgtggact gcgaggcggg cagctaccgg caaaacccag acgacatcgc ctgcaccttt 1620
tgtggccagg atgagtggtc cccggagcga agcacacgct gcttccgccg caggtctcgg 1680
ttcctggcat ggggcgagcc ggctgtgctg ctgctgctcc tgctgctgag cctggcgctg 1740
ggccttgtgc tggctgcttt ggggctgttc gttcaccatc gggacagccc actggttcag 1800
gcctcggggg ggcccctggc ctgctttggc ctggtgtgcc tgggcctggt ctgcctcagc 1860
gtcctcctgt tccctggcca gcccagccct gcccgatgcc tggcccagca gcccttgtcc 1920
cacctcccgc tcacgggctg cctgagcaca ctcttcctgc aggcggccga gatcttcgtg 1980
gagtcagaac tgcctctgag ctgggcagac cggctgagtg gctgcctgcg ggggccctgg 2040
gcctggctgg tggtgctgct ggccatgctg gtggaggtcg cactgtgcac ctggtacctg 2100
gtggccttcc cgccggaggt ggtgacggac tggcacatgc tgcccacgga ggcgctggtg 2160
cactgccgca cacgctcctg ggtcagcttc ggcctagcgc acgccaccaa tgccacgctg 2220
gcctttctct gcttcctggg cactttcctg gtgcggagcc agccgggccg ctacaaccgt 2280
gcccgtggcc tcacctttgc catgctggcc tacttcatca cctgggtctc ctttgtgccc 2340
ctcctggcca atgtgcaggt ggtcctcagg cccgccgtgc agatgggcgc cctcctgctc 2400
tgtgtcctgg gcatcctggc tgccttccac ctgcccaggt gttacctgct catgcggcag 2460
ccagggctca acacccccga gttcttcctg ggaggggggcc ctggggatgc ccaaggccag 2520
aatgacggga acacaggaaa tcaggggaaa catgagtga 2559

<210> 5
<211> 852
<212> PRT
<213> Homo sapiens

<400> 5
Met Leu Gly Pro Ala Val Leu Gly Leu Ser Leu Trp Ala Leu Leu His
1 5 10 15

Pro Gly Thr Gly Ala Pro Leu Cys Leu Ser Gln Gln Leu Arg Met Lys
20 25 30

Gly Asp Tyr Val Leu Gly Gly Leu Phe Pro Leu Gly Glu Ala Glu Glu
35 40 45

Ala Gly Leu Arg Ser Arg Thr Arg Pro Ser Ser Pro Val Cys Thr Arg
50 55 60

Phe Ser Ser Asn Gly Leu Leu Trp Ala Leu Ala Met Lys Met Ala Val
65 70 75 80

Glu Glu Ile Asn Asn Lys Ser Asp Leu Leu Pro Gly Leu Arg Leu Gly
85 90 95

```
Tyr Asp Leu Phe Asp Thr Cys Ser Glu Pro Val Val Ala Met Lys Pro
            100                 105                 110

Ser Leu Met Phe Leu Ala Lys Ala Gly Ser Arg Asp Ile Ala Ala Tyr
        115                 120                 125

Cys Asn Tyr Thr Gln Tyr Gln Pro Arg Val Leu Ala Val Ile Gly Pro
    130                 135                 140

His Ser Ser Glu Leu Ala Met Val Thr Gly Lys Phe Phe Ser Phe Phe
145                 150                 155                 160

Leu Met Pro Gln Val Ser Tyr Gly Ala Ser Met Glu Leu Leu Ser Ala
                165                 170                 175

Arg Glu Thr Phe Pro Ser Phe Phe Arg Thr Val Pro Ser Asp Arg Val
            180                 185                 190

Gln Leu Thr Ala Ala Ala Glu Leu Leu Gln Glu Phe Gly Trp Asn Trp
        195                 200                 205

Val Ala Ala Leu Gly Ser Asp Asp Glu Tyr Gly Arg Gln Gly Leu Ser
    210                 215                 220

Ile Phe Ser Ala Leu Ala Ser Ala Arg Gly Ile Cys Ile Ala His Glu
225                 230                 235                 240

Gly Leu Val Pro Leu Pro Arg Ala Asp Asp Ser Arg Leu Gly Lys Val
                245                 250                 255

Gln Asp Val Leu His Gln Val Asn Gln Ser Ser Val Gln Val Val Leu
            260                 265                 270

Leu Phe Ala Ser Val His Ala Ala His Ala Leu Phe Asn Tyr Ser Ile
        275                 280                 285

Ser Ser Arg Leu Ser Pro Lys Val Trp Val Ala Ser Glu Ala Trp Leu
    290                 295                 300

Thr Ser Asp Leu Val Met Gly Leu Pro Gly Met Ala Gln Met Gly Thr
305                 310                 315                 320

Val Leu Gly Phe Leu Gln Arg Gly Ala Gln Leu His Glu Phe Pro Gln
                325                 330                 335

Tyr Val Lys Thr His Leu Ala Leu Ala Thr Asp Pro Ala Phe Cys Ser
            340                 345                 350
```

```
Ala Leu Gly Glu Arg Glu Gln Gly Leu Glu Glu Asp Val Val Gly Gln
        355                 360                 365

Arg Cys Pro Gln Cys Asp Cys Ile Thr Leu Gln Asn Val Ser Ala Gly
    370                 375                 380

Leu Asn His His Gln Thr Phe Ser Val Tyr Ala Ala Val Tyr Ser Val
385                 390                 395                 400

Ala Gln Ala Leu His Asn Thr Leu Gln Cys Asn Ala Ser Gly Cys Pro
                405                 410                 415

Ala Gln Asp Pro Val Lys Pro Trp Gln Leu Leu Glu Asn Met Tyr Asn
            420                 425                 430

Leu Thr Phe His Val Gly Gly Leu Pro Leu Arg Phe Asp Ser Ser Gly
        435                 440                 445

Asn Val Asp Met Glu Tyr Asp Leu Lys Leu Trp Val Trp Gln Gly Ser
    450                 455                 460

Val Pro Arg Leu His Asp Val Gly Arg Phe Asn Gly Ser Leu Arg Thr
465                 470                 475                 480

Glu Arg Leu Lys Ile Arg Trp His Thr Ser Asp Asn Gln Lys Pro Val
                485                 490                 495

Ser Arg Cys Ser Arg Gln Cys Gln Glu Gly Gln Val Arg Arg Val Lys
            500                 505                 510

Gly Phe His Ser Cys Cys Tyr Asp Cys Val Asp Cys Glu Ala Gly Ser
        515                 520                 525

Tyr Arg Gln Asn Pro Asp Asp Ile Ala Cys Thr Phe Cys Gly Gln Asp
    530                 535                 540

Glu Trp Ser Pro Glu Arg Ser Thr Arg Cys Phe Arg Arg Arg Ser Arg
545                 550                 555                 560

Phe Leu Ala Trp Gly Glu Pro Ala Val Leu Leu Leu Leu Leu Leu Leu
                565                 570                 575

Ser Leu Ala Leu Gly Leu Val Leu Ala Ala Leu Gly Leu Phe Val His
            580                 585                 590

His Arg Asp Ser Pro Leu Val Gln Ala Ser Gly Gly Pro Leu Ala Cys
        595                 600                 605
```

```
Phe Gly Leu Val Cys Leu Gly Leu Val Cys Leu Ser Val Leu Leu Phe
    610                 615                 620

Pro Gly Gln Pro Ser Pro Ala Arg Cys Leu Ala Gln Gln Pro Leu Ser
625                 630                 635                 640

His Leu Pro Leu Thr Gly Cys Leu Ser Thr Leu Phe Leu Gln Ala Ala
                645                 650                 655

Glu Ile Phe Val Glu Ser Glu Leu Pro Leu Ser Trp Ala Asp Arg Leu
            660                 665                 670

Ser Gly Cys Leu Arg Gly Pro Trp Ala Trp Leu Val Val Leu Leu Ala
        675                 680                 685

Met Leu Val Glu Val Ala Leu Cys Thr Trp Tyr Leu Val Ala Phe Pro
    690                 695                 700

Pro Glu Val Val Thr Asp Trp His Met Leu Pro Thr Glu Ala Leu Val
705                 710                 715                 720

His Cys Arg Thr Arg Ser Trp Val Ser Phe Gly Leu Ala His Ala Thr
                725                 730                 735

Asn Ala Thr Leu Ala Phe Leu Cys Phe Leu Gly Thr Phe Leu Val Arg
            740                 745                 750

Ser Gln Pro Gly Arg Tyr Asn Arg Ala Arg Gly Leu Thr Phe Ala Met
        755                 760                 765

Leu Ala Tyr Phe Ile Thr Trp Val Ser Phe Val Pro Leu Leu Ala Asn
    770                 775                 780

Val Gln Val Val Leu Arg Pro Ala Val Gln Met Gly Ala Leu Leu Leu
785                 790                 795                 800

Cys Val Leu Gly Ile Leu Ala Ala Phe His Leu Pro Arg Cys Tyr Leu
                805                 810                 815

Leu Met Arg Gln Pro Gly Leu Asn Thr Pro Glu Phe Phe Leu Gly Gly
            820                 825                 830

Gly Pro Gly Asp Ala Gln Gly Gln Asn Asp Gly Asn Thr Gly Asn Gln
        835                 840                 845

Gly Lys His Glu
    850
```

<210> 6
<211> 20
<212> DNA
<213> Mouse

<400> 6
cactagagct gccaccttcc 20

<210> 7
<211> 20
<212> DNA
<213> Mouse

<400> 7
ccctcagcac cacttttgt 20

<210> 8
<211> 20
<212> DNA
<213> Mouse

<400> 8
acaaaaagtg gtgctgaggg 20

<210> 9
<211> 20
<212> DNA
<213> Mouse

<400> 9
caggagaccc aaaggatcaa 20

<210> 10
<211> 20
<212> DNA
<213> Mouse

<400> 10
gcttcagaaa atcgaggcac 20

<210> 11
<211> 20

<212> DNA
<213> Mouse

<400> 11
gcatgggcta tgataggtgg 20

<210> 12
<211> 16
<212> DNA
<213> Mouse

<400> 12
tgttgatccc acagcg 16

<210> 13
<211> 20
<212> DNA
<213> Mouse

<400> 13
caggaaatgt ccacttctgc 20

<210> 14
<211> 18
<212> DNA
<213> Mouse

<400> 14
tctatcttgc atccagcc 18

<210> 15
<211> 16
<212> DNA
<213> Mouse

<400> 15
gtgctgtgac tgtgcg 16

<210> 16
<211> 18
<212> DNA
<213> Mouse

<400> 16
cgcagcattt atttggag 18

<210> 17
<211> 19
<212> DNA
<213> Mouse

<400> 17
ccgacccttt aggagacac 19

<210> 18
<211> 20
<212> DNA
<213> Mouse

<400> 18
tgtgacttcc tcttccccac 20

<210> 19
<211> 20
<212> DNA
<213> Mouse

<400> 19
tgagccactc cagatgtcag 20

<210> 20
<211> 20
<212> DNA
<213> Mouse

<400> 20
ccaacgtgca gtcaagaaaa 20

<210> 21
<211> 20
<212> DNA
<213> Mouse

<400> 21
ccaacgtgca gtcaagaaaa 20

<210> 22
<211> 20
<212> DNA
<213> Mouse

<400> 22
cgagagacaa agtggtgctg 20

<210> 23
<211> 20
<212> DNA
<213> Mouse

<400> 23
ttatgaaggc cctcaccaac 20

<210> 24
<211> 20
<212> DNA
<213> Mouse

<400> 24
ccagctccta gaattgcctg 20

<210> 25
<211> 20
<212> DNA
<213> Mouse

<400> 25
gcagtctccc gaaacaagtc 20

<210> 26
<211> 20
<212> DNA
<213> Mouse

<400> 26
atagaggaat gggtgcgatg 20

<210> 27
<211> 20

<212> DNA
<213> Mouse

<400> 27
taccaggagg ggtcagtcag 20

<210> 28
<211> 20
<212> DNA
<213> Mouse

<400> 28
tacaagcgag ctgaccaatg 20

<210> 29
<211> 20
<212> DNA
<213> Mouse

<400> 29
ccaatcagct cgagttagcc 20

<210> 30
<211> 20
<212> DNA
<213> Mouse

<400> 30
tgccattgtg gatgttcact 20

<210> 31
<211> 20
<212> DNA
<213> Mouse

<400> 31
gagtccgagg tcggtcaata 20

<210> 32
<211> 20
<212> DNA
<213> Mouse

<400> 32
gctggcttct gtaggtcagg 20

<210> 33
<211> 20
<212> DNA
<213> Mouse

<400> 33
tatgagggtc aagggtcagg 20

<210> 34
<211> 20
<212> DNA
<213> Mouse

<400> 34
cgctttggtg agaactagcc 20

<210> 35
<211> 20
<212> DNA
<213> Mouse

<400> 35
catgtggagt tgtgggagtg 20

<210> 36
<211> 20
<212> DNA
<213> Mouse

<400> 36
aatgggcaga agacagatgg 20

<210> 37
<211> 20
<212> DNA
<213> Mouse

<400> 37
tatcagggtc tgtgaagccc 20

<210> 38
<211> 20
<212> DNA
<213> Mouse

<400> 38
atacaggacc ctttaccccg 20

<210> 39
<211> 20
<212> DNA
<213> Mouse

<400> 39
cagtgtttct aggtccccca 20

<210> 40
<211> 20
<212> DNA
<213> Mouse

<400> 40
gcctctgtct gccatctctc 20

<210> 41
<211> 20
<212> DNA
<213> Mouse

<400> 41
ataatgttac ctgcaggcgg 20

<210> 42
<211> 20
<212> DNA
<213> Mouse

<400> 42
ctggaaacac ccatgtcctc 20

<210> 43
<211> 20

<212> DNA
<213> Mouse

<400> 43
cgggcacatg gacactttta 20

<210> 44
<211> 20
<212> DNA
<213> Mouse

<400> 44
gagcatgaag tgcaaggtga 20

<210> 45
<211> 20
<212> DNA
<213> Mouse

<400> 45
cgtaggtggc acagttgaga 20

<210> 46
<211> 20
<212> DNA
<213> Mouse

<400> 46
gctgttagtg aggtcagggc 20

<210> 47
<211> 20
<212> DNA
<213> Mouse

<400> 47
cgtaggtggc acagttgaga 20

<210> 48
<211> 20
<212> DNA
<213> Mouse

<400> 48
gagcatgaag tgcaaggtga 20

<210> 49
<211> 20
<212> DNA
<213> Mouse

<400> 49
tcattttcct agcctcggtg 20

<210> 50
<211> 22
<212> DNA
<213> Mouse

<400> 50
tctaagaaga tgatgcagac cc 22

<210> 51
<211> 20
<212> DNA
<213> Mouse

<400> 51
tgtccttcag ggatagtgcc 20

<210> 52
<211> 20
<212> DNA
<213> Mouse

<400> 52
ggcttcagcc tcaagttctg 20

<210> 53
<211> 20
<212> DNA
<213> Mouse

<400> 53
aaaacaacca agttgccctg 20

<210> 54
<211> 20
<212> DNA
<213> Mouse

<400> 54
ggcactgaaa tgacctggat 20

<210> 55
<211> 20
<212> DNA
<213> Mouse

<400> 55
aacaattcaa gcaacctcgg 20

<210> 56
<211> 20
<212> DNA
<213> Mouse

<400> 56
ctgttccttc ccagactcca 20

<210> 57
<211> 20
<212> DNA
<213> Mouse

<400> 57
ttcagtcacg caaacctgag 20

<210> 58
<211> 20
<212> DNA
<213> Mouse

<400> 58
gcccaggact ttgtcactgt 20

<210> 59
<211> 20

<212> DNA
<213> Mouse

<400> 59
ggtaacctgc agctccactc 20

<210> 60
<211> 20
<212> DNA
<213> Mouse

<400> 60
gggacatgct cttggttcat 20

<210> 61
<211> 20
<212> DNA
<213> Mouse

<400> 61
gaacaaagcc gggtgattta 20

<210> 62
<211> 20
<212> DNA
<213> Mouse

<400> 62
gccctcagtt ctcctagcct 20

<210> 63
<211> 20
<212> DNA
<213> Mouse

<400> 63
ggcagagaag actggtggag 20

<210> 64
<211> 20
<212> DNA
<213> Mouse

<400> 64
cccagactta gcgtctcagg 20

<210> 65
<211> 20
<212> DNA
<213> Mouse

<400> 65
agcagagacc tttggactcg 20

<210> 66
<211> 20
<212> DNA
<213> Mouse

<400> 66
gaaggctgag tgagtcccag 20

<210> 67
<211> 20
<212> DNA
<213> Mouse

<400> 67
ttgcacgagg agaaggtttt 20

<210> 68
<211> 20
<212> DNA
<213> Mouse

<400> 68
gatgccaacg agacctgaat 20

<210> 69
<211> 20
<212> DNA
<213> Mouse

<400> 69
agaagccaaa accctcacct 20

<210> 70
<211> 20
<212> DNA
<213> Mouse

<400> 70
aaaaagccct gcaagaactt 20

<210> 71
<211> 20
<212> DNA
<213> Mouse

<400> 71
attcaggtct cgttggcatc 20

<210> 72
<211> 20
<212> DNA
<213> Mouse

<400> 72
tgtccgcagt gtggaaacta 20

<210> 73
<211> 20
<212> DNA
<213> Mouse

<400> 73
atgtccaggg tagagagccc 20

<210> 74
<211> 20
<212> DNA
<213> Mouse

<400> 74
ggagttctcc taccctggct 20

<210> 75
<211> 20

<212> DNA
<213> Mouse

<400> 75
gaggctctga gcagtgtcaa 20

<210> 76
<211> 14
<212> DNA
<213> Mouse

<400> 76
gcgatgttgt tgcg 14

<210> 77
<211> 18
<212> DNA
<213> Mouse

<400> 77
cagtgtcttt ccacattt 18

<210> 78
<211> 27
<212> DNA
<213> Mouse

<400> 78
aggcatattg tataataaat ttgtagt 27

<210> 79
<211> 19
<212> DNA
<213> Mouse

<400> 79
ccggatgact ctacttgac 19

<210> 80
<211> 20
<212> DNA
<213> Mouse

<400> 80
gctgtttatg gggtcgagaa 20

<210> 81
<211> 20
<212> DNA
<213> Mouse

<400> 81
aatttctgaa gcaggggggat 20

<210> 82
<211> 20
<212> DNA
<213> Mouse

<400> 82
tccccctgct tcagaaatta 20

<210> 83
<211> 20
<212> DNA
<213> Mouse

<400> 83
agggggatga ttgtgagtga 20

<210> 84
<211> 27
<212> DNA
<213> Mouse

<400> 84
cttctttaat caatctctgt ctctgtg 27

<210> 85
<211> 20
<212> DNA
<213> Mouse

<400> 85
gggcacatat gaacctcctg 20

<210> 86
<211> 20
<212> DNA
<213> Mouse

<400> 86
ccaaactctt agcttcttca 20

<210> 87
<211> 21
<212> DNA
<213> Mouse

<400> 87
acacagaaga cactgaagaa c 21

<210> 88
<211> 22
<212> DNA
<213> Mouse

<400> 88
cagttgttag aagcaggatc cc 22

<210> 89
<211> 23
<212> DNA
<213> Mouse

<400> 89
aggtgcatat acctgggata ctc 23

<210> 90
<211> 21
<212> DNA
<213> Mouse

<400> 90
agagtttggt ctcttcccct g 21

<210> 91
<211> 23

<212> DNA
<213> Mouse

<400> 91
tatccaacac atttatgtct gcg 23

<210> 92
<211> 20
<212> DNA
<213> Mouse

<400> 92
gccagtgtgc tgaaagactg 20

<210> 93
<211> 20
<212> DNA
<213> Mouse

<400> 93
agggacctgg agacatcctt 20

<210> 94
<211> 23
<212> DNA
<213> Mouse

<400> 94
ctgtaggctg cttttatctt ttg 23

<210> 95
<211> 20
<212> DNA
<213> Mouse

<400> 95
tgccccttca gcacatgcca 20

<210> 96
<211> 23
<212> DNA
<213> Mouse

<400> 96
tgcagtgtga catgtgcata gat 23

<210> 97
<211> 21
<212> DNA
<213> Mouse

<400> 97
ggaaagccag gctacgcaga a 21

<210> 98
<211> 23
<212> DNA
<213> Mouse

<400> 98
ctgtaggctg cttttatctt ttg 23

<210> 99
<211> 20
<212> DNA
<213> Mouse

<400> 99
tgccccttca gcacatgcca 20

<210> 100
<211> 22
<212> DNA
<213> Mouse

<400> 100
tagtgtggtt cctgactaac ct 22

<210> 101
<211> 22
<212> DNA
<213> Mouse

<400> 101
cggtctacat agtgagtgat tc 22

<210> 102
<211> 22
<212> DNA
<213> Mouse

<400> 102
aaaagcatcc tgcatccttc tg 22

<210> 103
<211> 22
<212> DNA
<213> Mouse

<400> 103
gggttataca gagaaaccct gt 22

<210> 104
<211> 20
<212> DNA
<213> Mouse

<400> 104
ttccaagctc acacatcagc 20

<210> 105
<211> 20
<212> DNA
<213> Mouse

<400> 105
gtgctgctct gcattgagtg 20

<210> 106
<211> 20
<212> DNA
<213> Mouse

<400> 106
gacagtgtgg gagaatccgt 20

<210> 107
<211> 20

<212> DNA
<213> Mouse

<400> 107
cccaaggcat aggtcacaat 20

<210> 108
<211> 20
<212> DNA
<213> Mouse

<400> 108
attgtgacct atgccttggg 20

<210> 109
<211> 20
<212> DNA
<213> Mouse

<400> 109
cgaaggaccg tcatctgagt 20

<210> 110
<211> 20
<212> DNA
<213> Mouse

<400> 110
ggctttgatg tgaaaaaggc 20

<210> 111
<211> 20
<212> DNA
<213> Mouse

<400> 111
agctcctcat cgctcatgtt 20

<210> 112
<211> 20
<212> DNA
<213> Mouse

<400> 112
tggaacatct ctgtcggaag 20

<210> 113
<211> 20
<212> DNA
<213> Mouse

<400> 113
ggctctcatt gccaccttta 20

<210> 114
<211> 20
<212> DNA
<213> Mouse

<400> 114
ccagagaaca ggagacctgc 20

<210> 115
<211> 20
<212> DNA
<213> Mouse

<400> 115
gtgctggata cactggcaga 20

<210> 116
<211> 20
<212> DNA
<213> Mouse

<400> 116
gcgagacgag tgggtagttc 20

<210> 117
<211> 20
<212> DNA
<213> Mouse

<400> 117
acactgaaac ctcgcttgct 20

<210> 118
<211> 20
<212> DNA
<213> Mouse

<400> 118
agcaagcgag gtttcagtgt 20

<210> 119
<211> 20
<212> DNA
<213> Mouse

<400> 119
acggggcttg atccttttat 20

<210> 120
<211> 25
<212> DNA
<213> Mouse

<400> 120
aagttcatgg gcctcaccac ctgtc 25

<210> 121
<211> 22
<212> DNA
<213> Mouse

<400> 121
tactagctac ccttcacata cc 22

<210> 122
<211> 21
<212> DNA
<213> Mouse

<400> 122
acctagccac tgtctcagtc t 21

<210> 123
<211> 21

<212> DNA
<213> Mouse

<400> 123
acagaagcag catttacaca g 21

<210> 124
<211> 20
<212> DNA
<213> Mouse

<400> 124
tgggacagct tcctcaagat 20

<210> 125
<211> 20
<212> DNA
<213> Mouse

<400> 125
aatgggaatt gtgctcttgg 20

<210> 126
<211> 20
<212> DNA
<213> Mouse

<400> 126
gggcatctgg caaagattta 20

<210> 127
<211> 20
<212> DNA
<213> Mouse

<400> 127
agataacctg tgtgtcccgc 20

<210> 128
<211> 20
<212> DNA
<213> Mouse

<400> 128
gatgtccgag aagggatgtg 20

<210> 129
<211> 20
<212> DNA
<213> Mouse

<400> 129
tgtcagcttt gagtgcatcc 20

<210> 130
<211> 20
<212> DNA
<213> Mouse

<400> 130
acatgcaggc tgtttgacct 20

<210> 131
<211> 20
<212> DNA
<213> Mouse

<400> 131
tgtcagcttt gagtgcatcc 20

<210> 132
<211> 20
<212> DNA
<213> Mouse

<400> 132
gtgctctgca gacaaaccaa 20

<210> 133
<211> 20
<212> DNA
<213> Mouse

<400> 133
gagccatttt gacccttaaa 20

<210> 134
<211> 20
<212> DNA
<213> Mouse

<400> 134
tttcagggtc aaaatggctc 20

<210> 135
<211> 17
<212> DNA
<213> Mouse

<400> 135
tcgacagcaa ctgtgcg 17

<210> 136
<211> 20
<212> DNA
<213> Mouse

<400> 136
ggtgagagtg gggagatgaa 20

<210> 137
<211> 20
<212> DNA
<213> Mouse

<400> 137
cccgggtgag tttaagaacc 20

<210> 138
<211> 20
<212> DNA
<213> Mouse

<400> 138
ggtgagagtg gggagatgaa 20

<210> 139
<211> 20

<212> DNA
<213> Mouse

<400> 139
aggttaggcc caatttcctg 20

<210> 140
<211> 20
<212> DNA
<213> Mouse

<400> 140
ccagggttgc tgtactgaga 20

<210> 141
<211> 20
<212> DNA
<213> Mouse

<400> 141
caggttaggc ccaatttcct 20

<210> 142
<211> 20
<212> DNA
<213> Mouse

<400> 142
ggtcagagtc cttccttccc 20

<210> 143
<211> 20
<212> DNA
<213> Mouse

<400> 143
tccaacttca caggaaaccc 20

<210> 144
<211> 20
<212> DNA
<213> Mouse

<400> 144
tttcctgtga agttggaggg 20

<210> 145
<211> 20
<212> DNA
<213> Mouse

<400> 145
cacccatatg gcaaacatca 20

<210> 146
<211> 20
<212> DNA
<213> Mouse

<400> 146
ggtcagagtc cttccttccc 20

<210> 147
<211> 20
<212> DNA
<213> Mouse

<400> 147
tccaacttca caggaaaccc 20

<210> 148
<211> 20
<212> DNA
<213> Mouse

<400> 148
tgatgtttgc catatgggtg 20

<210> 149
<211> 20
<212> DNA
<213> Mouse

<400> 149
gcttgctgct tccgatatgt 20

<210> 150
<211> 19
<212> DNA
<213> Mouse

<400> 150
ggaaaaggga gtcgccata 19

<210> 151
<211> 20
<212> DNA
<213> Mouse

<400> 151
gagccgccta actctcacac 20

<210> 152
<211> 19
<212> DNA
<213> Mouse

<400> 152
aggggataac ctgcatagg 19

<210> 153
<211> 20
<212> DNA
<213> Mouse

<400> 153
acaaaattgc tcatttgccc 20

<210> 154
<211> 20
<212> DNA
<213> Mouse

<400> 154
ccatccccac tagccagata 20

<210> 155
<211> 20

<212> DNA
<213> Mouse

<400> 155
gtccccttg tcacagcaag 20

<210> 156
<211> 20
<212> DNA
<213> Mouse

<400> 156
tgagcacagg atagctccac 20

<210> 157
<211> 20
<212> DNA
<213> Mouse

<400> 157
aaaagaacac ctgtttgggg 20

<210> 158
<211> 19
<212> DNA
<213> Mouse

<400> 158
taaacctcgg ctgtgtgag 19

<210> 159
<211> 20
<212> DNA
<213> Mouse

<400> 159
ccctcagtga cttcctgtga 20

<210> 160
<211> 20
<212> DNA
<213> Mouse

<400> 160
caaaaccaca tggttaccga 20

<210> 161
<211> 20
<212> DNA
<213> Mouse

<400> 161
gccctattgc caaatgactt 20

<210> 162
<211> 20
<212> DNA
<213> Mouse

<400> 162
ggcagaaagg aatcagaagc 20

<210> 163
<211> 20
<212> DNA
<213> Mouse

<400> 163
cacattagcc attgtcctgg 20

<210> 164
<211> 20
<212> DNA
<213> Mouse

<400> 164
tcctttatgt ccaacagcca 20

<210> 165
<211> 20
<212> DNA
<213> Mouse

<400> 165
catggtctgt gatgtgacca 20

<210> 166
<211> 20
<212> DNA
<213> Mouse

<400> 166
ataccettgg tgagagcagg 20

<210> 167
<211> 20
<212> DNA
<213> Mouse

<400> 167
gctgtcaaat gagaaaggca 20

<210> 168
<211> 20
<212> DNA
<213> Mouse

<400> 168
tatttcatgc tgggaccaaa 20

<210> 169
<211> 20
<212> DNA
<213> Mouse

<400> 169
agagaaaaac agtgggggtg 20

<210> 170
<211> 20
<212> DNA
<213> Mouse

<400> 170
cgggtcctct cttcaccata 20

<210> 171
<211> 20

<212> DNA
<213> Mouse

<400> 171
ctacatttcc ctgagctgcc 20

<210> 172
<211> 20
<212> DNA
<213> Mouse

<400> 172
gttgaccatg tcggtaaccc 20

<210> 173
<211> 20
<212> DNA
<213> Mouse

<400> 173
ccacctcacg gaaactgaat 20

<210> 174
<211> 20
<212> DNA
<213> Mouse

<400> 174
ggtgtttggc tcacaaacct 20

<210> 175
<211> 20
<212> DNA
<213> Mouse

<400> 175
gatgcacaca caaaaatccg 20

<210> 176
<211> 20
<212> DNA
<213> Mouse

<400> 176
atcacccacc agaacgaaaa 20

<210> 177
<211> 20
<212> DNA
<213> Mouse

<400> 177
accctccagg agtaggtgct 20

<210> 178
<211> 20
<212> DNA
<213> Mouse

<400> 178
gatgagacag tgggcaaggt 20

<210> 179
<211> 20
<212> DNA
<213> Mouse

<400> 179
ttgtcaatag caccaagcca 20

<210> 180
<211> 20
<212> DNA
<213> Mouse

<400> 180
gccttaatag ccccttgtt 20

<210> 181
<211> 20
<212> DNA
<213> Mouse

<400> 181
gcactcagca ttgcacagat 20

<210> 182
<211> 20
<212> DNA
<213> Mouse

<400> 182
ggacggacaa ttctggaaaa 20

<210> 183
<211> 20
<212> DNA
<213> Mouse

<400> 183
ctatcacacc tccgatgcct 20

<210> 184
<211> 20
<212> DNA
<213> Mouse

<400> 184
caagctggta gaatccccaa 20

<210> 185
<211> 20
<212> DNA
<213> Mouse

<400> 185
tctttggaga agcagaccgt 20

<210> 186
<211> 20
<212> DNA
<213> Mouse

<400> 186
tacagcatat gcatgccagg 20

<210> 187
<211> 20

<212> DNA
<213> Mouse

<400> 187
attcctcagg gcattacacg 20

<210> 188
<211> 20
<212> DNA
<213> Mouse

<400> 188
gcaatctctt gtgtccaggc 20

<210> 189
<211> 20
<212> DNA
<213> Mouse

<400> 189
attcctcagg gcattacacg 20

<210> 190
<211> 20
<212> DNA
<213> Mouse

<400> 190
tacagcatat gcatgccagg 20

<210> 191
<211> 20
<212> DNA
<213> Mouse

<400> 191
ggcctggaca caagagattg 20

<210> 192
<211> 20
<212> DNA
<213> Mouse

<400> 192
aagtgggtgg acagtgaagg 20

<210> 193
<211> 20
<212> DNA
<213> Mouse

<400> 193
cagcttcctc catcttctgg 20

<210> 194
<211> 20
<212> DNA
<213> Mouse

<400> 194
agagcctcca gtagatggca 20

<210> 195
<211> 20
<212> DNA
<213> Mouse

<400> 195
tcgtggacaa gctccttctt 20

<210> 196
<211> 20
<212> DNA
<213> Mouse

<400> 196
catcgagtat gtcaatggcg 20

<210> 197
<211> 20
<212> DNA
<213> Mouse

<400> 197
ttgtccagtt ttaggtcccg 20

<210> 198
<211> 20
<212> DNA
<213> Mouse

<400> 198
cagactgggt tttccgacat 20

<210> 199
<211> 20
<212> DNA
<213> Mouse

<400> 199
gtcaaagttg tccaggccat 20

<210> 200
<211> 18
<212> DNA
<213> Mouse

<400> 200
aggacggacc ccaagatg 18

<210> 201
<211> 20
<212> DNA
<213> Mouse

<400> 201
tgtctcgcac ttcctcacag 20

<210> 202
<211> 20
<212> DNA
<213> Mouse

<400> 202
ccagaagatg gaggaagctg 20

<210> 203
<211> 20

<212> DNA
<213> Mouse

<400> 203
tctactggag gctcttggga 20

<210> 204
<211> 20
<212> DNA
<213> Mouse

<400> 204
gaaaaacgac cagatttacg 20

<210> 205
<211> 20
<212> DNA
<213> Mouse

<400> 205
gatctcagca gcatagaacc 20

<210> 206
<211> 20
<212> DNA
<213> Mouse

<400> 206
acacattaag ctgacggact 20

<210> 207
<211> 20
<212> DNA
<213> Mouse

<400> 207
caaacataag gacacccagt 20

<210> 208
<211> 20
<212> DNA
<213> Mouse

<400> 208
actgggtgtc cttatgtttg 20

<210> 209
<211> 20
<212> DNA
<213> Mouse

<400> 209
cctctctttg ggatccttat 20

<210> 210
<211> 20
<212> DNA
<213> Mouse

<400> 210
gtcataaaga ggatcgacca 20

<210> 211
<211> 20
<212> DNA
<213> Mouse

<400> 211
gctctgtcta gaagtgcctg 20

<210> 212
<211> 18
<212> DNA
<213> Mouse

<400> 212
accaagaccg aagagggg 18

<210> 213
<211> 22
<212> DNA
<213> Mouse

<400> 213
ggcattacac gctaactttt cc 22

<210> 214
<211> 20
<212> DNA
<213> Mouse

<400> 214
agtgccacca acctggtaag 20

<210> 215
<211> 18
<212> DNA
<213> Mouse

<400> 215
aagtgcctgc agggatgc 18

<210> 216
<211> 20
<212> DNA
<213> Mouse

<400> 216
tgctttggtg agcaatgttt 20

<210> 217
<211> 20
<212> DNA
<213> Mouse

<400> 217
agggacaccc ttaccaggtt 20

<210> 218
<211> 20
<212> DNA
<213> Mouse

<400> 218
ctgatgcttt ggtgagcaat 20

<210> 219
<211> 19

<212> DNA
<213> Mouse

<400> 219
gggacaccct taccaggtt 19

<210> 220
<211> 20
<212> DNA
<213> Mouse

<400> 220
acaggacaaa tgctgggttg 20

<210> 221
<211> 20
<212> DNA
<213> Mouse

<400> 221
gtggtaaaga acgcttggct 20

<210> 222
<211> 24
<212> DNA
<213> Mouse

<400> 222
ggtatctcac ttggtaggaa cctc 24

<210> 223
<211> 17
<212> DNA
<213> Mouse

<400> 223
aagaacgctt ggctggc 17

<210> 224
<211> 20
<212> DNA
<213> Mouse

<400> 224
gccgatcctg gtgatgtact 20

<210> 225
<211> 20
<212> DNA
<213> Mouse

<400> 225
acaatggctc aaaaccgttc 20

<210> 226
<211> 20
<212> DNA
<213> Mouse

<400> 226
gccttgggaa tttaccacct 20

<210> 227
<211> 20
<212> DNA
<213> Mouse

<400> 227
agtacatcac caggatcggc 20

<210> 228
<211> 20
<212> DNA
<213> Mouse

<400> 228
taaaaggcca tgcgataagc 20

<210> 229
<211> 20
<212> DNA
<213> Mouse

<400> 229
agagctctgt ggggttctca 20

<210> 230
<211> 20
<212> DNA
<213> Mouse

<400> 230
gaaggggaca gtgttggaga 20

<210> 231
<211> 20
<212> DNA
<213> Mouse

<400> 231
tccatcaagg aaggatccac 20

<210> 232
<211> 19
<212> DNA
<213> Mouse

<400> 232
ggtgggtaat gattggact 19

<210> 233
<211> 19
<212> DNA
<213> Mouse

<400> 233
tgacgtggag ggaactgcc 19

<210> 234
<211> 20
<212> DNA
<213> Mouse

<400> 234
tgagatctgg tgccctctct 20

<210> 235
<211> 20

<212> DNA
<213> Mouse

<400> 235
gcctgatcta ggctggaaaa 20

<210> 236
<211> 20
<212> DNA
<213> Mouse

<400> 236
aggcagaaag cagacaagga 20

<210> 237
<211> 20
<212> DNA
<213> Mouse

<400> 237
cgacagcact tgtgaccact 20

<210> 238
<211> 20
<212> DNA
<213> Mouse

<400> 238
ctgcagatgt agaccaggca 20

<210> 239
<211> 20
<212> DNA
<213> Mouse

<400> 239
ctgtggtgga ttggacagtg 20

<210> 240
<211> 20
<212> DNA
<213> Mouse

<400> 240
ttgcctaaca ctcccaaacc 20

<210> 241
<211> 20
<212> DNA
<213> Mouse

<400> 241
tattaggagc accaccaggc 20

<210> 242
<211> 20
<212> DNA
<213> Mouse

<400> 242
acctgtcttg tgggtggaag 20

<210> 243
<211> 20
<212> DNA
<213> Mouse

<400> 243
ctgtggtgga ttggacagtg 20

<210> 244
<211> 20
<212> DNA
<213> Mouse

<400> 244
gtggcttggt gctattgaca 20

<210> 245
<211> 20
<212> DNA
<213> Mouse

<400> 245
ggggctatta aggccatttt 20

<210> 246
<211> 21
<212> DNA
<213> Mouse

<400> 246
caattgagga atggctacca a 21

<210> 247
<211> 20
<212> DNA
<213> Mouse

<400> 247
tggcttcatg tccattgtgt 20

<210> 248
<211> 22
<212> DNA
<213> Mouse

<400> 248
cagaaccaca aaggtaaatt gc 22

<210> 249
<211> 21
<212> DNA
<213> Mouse

<400> 249
tcatgtttgc tgtccagttt g 21

<210> 250
<211> 29
<212> DNA
<213> Homo sapiens

<400> 250
gccaccatgc tgggccctgc tgtcctggg 29

<210> 251
<211> 24

<212> DNA
<213> Homo sapiens

<400> 251
tcactcatgt ttcccctgat ttcc 24

<210> 252
<211> 20
<212> DNA
<213> Homo sapiens

<400> 252
ctgatttcct gtgttcccgt 20

<210> 253
<211> 20
<212> DNA
<213> Homo sapiens

<400> 253
catgctggcc tacttcatca 20

<210> 254
<211> 29
<212> DNA
<213> Homo sapiens

<400> 254
gccttgcagg tcagctacgg tgctagcat 29

<210> 255
<211> 24
<212> DNA
<213> Homo sapiens

<400> 255
tcactcatgt ttcccctgat ttcc 24

<210> 256
<211> 20
<212> DNA
<213> Homo sapiens

<400> 256
aggaagcaga gaaaggccag 20

<210> 257
<211> 20
<212> DNA
<213> Homo sapiens

<400> 257
tcagaactgc ctctgagctg 20

<210> 258
<211> 20
<212> DNA
<213> Homo sapiens

<400> 258
tcttcacgta ctgggggaac 20

<210> 259
<211> 20
<212> DNA
<213> Homo sapiens

<400> 259
actacagcat cagcagcagg 20

<210> 260
<211> 20
<212> DNA
<213> Homo sapiens

<400> 260
aagctgaaga acttcccggt 20

<210> 261
<211> 20
<212> DNA
<213> Homo sapiens

<400> 261
tgggctacga cctctttgat 20

<210> 262
<211> 20
<212> DNA
<213> Homo sapiens

<400> 262
atcttcaggc gctctgtcct 20

<210> 263
<211> 20
<212> DNA
<213> Homo sapiens

<400> 263
gtacgacctg aagctgtggg 20

<210> 264
<211> 19
<212> DNA
<213> Homo sapiens

<400> 264
atcttcaggc gctctgtcc 19

<210> 265
<211> 20
<212> DNA
<213> Homo sapiens

<400> 265
gtacgacctg aagctgtggg 20

<210> 266
<211> 19
<212> DNA
<213> Homo sapiens

<400> 266
atcttcaggc gctctgtcc 19

<210> 267
<211> 21

<212> DNA
<213> Homo sapiens

<400> 267
gagtacgacc tgaagctgtg g 21

<210> 268
<211> 19
<212> DNA
<213> Homo sapiens

<400> 268
atcttcaggc gctctgtcc 19

<210> 269
<211> 19
<212> DNA
<213> Homo sapiens

<400> 269
tacgacctga agctgtggg 19

<210> 270
<211> 19
<212> DNA
<213> Homo sapiens

<400> 270
atcttcaggc gctctgtcc 19

<210> 271
<211> 19
<212> DNA
<213> Homo sapiens

<400> 271
tacgacctga agctgtggg 19

<210> 272
<211> 18
<212> DNA
<213> Homo sapiens

<400> 272
gctgtcccga tggtgaac 18

<210> 273
<211> 19
<212> DNA
<213> Homo sapiens

<400> 273
acctttgtg gccaggatg 19

<210> 274
<211> 18
<212> DNA
<213> Homo sapiens

<400> 274
gctgtcccga tggtgaac 18

<210> 275
<211> 19
<212> DNA
<213> Homo sapiens

<400> 275
cacctttgt ggccaggat 19

<210> 276
<211> 18
<212> DNA
<213> Homo sapiens

<400> 276
gctgtcccga tggtgaac 18

<210> 277
<211> 18
<212> DNA
<213> Homo sapiens

<400> 277
cctttgtgg ccaggatg 18

<210> 278
<211> 18
<212> DNA
<213> Homo sapiens

<400> 278
cctgaaccag tgggctgt 18

<210> 279
<211> 19
<212> DNA
<213> Homo sapiens

<400> 279
acctttgtg gccaggatg 19

<210> 280
<211> 18
<212> DNA
<213> Homo sapiens

<400> 280
cctgaaccag tgggctgt 18

<210> 281
<211> 19
<212> DNA
<213> Homo sapiens

<400> 281
caccttttgt ggccaggat 19

<210> 282
<211> 20
<212> DNA
<213> Homo sapiens

<400> 282
tcatgtttcc cctgatttcc 20

<210> 283
<211> 20

<212> DNA
<213> Homo sapiens

<400> 283
catgctggcc tacttcatca 20

<210> 284
<211> 20
<212> DNA
<213> Homo sapiens

<400> 284
atgagcaggt aacacctggg 20

<210> 285
<211> 20
<212> DNA
<213> Homo sapiens

<400> 285
tcatcacctg ggtctccttt 20

<210> 286
<211> 20
<212> DNA
<213> Homo sapiens

<400> 286
atgagcaggt aacacctggg 20

<210> 287
<211> 20
<212> DNA
<213> Homo sapiens

<400> 287
ttcatcacct gggtctcctt 20

<210> 288
<211> 20
<212> DNA
<213> Mouse

<400> 288
tgggttgtgt tctctggttg 20

<210> 289
<211> 21
<212> DNA
<213> Mouse

<400> 289
ccttttttaca gtctgccagg t 21

<210> 290
<211> 20
<212> DNA
<213> Mouse

<400> 290
tgggttgtgt tctctggttg 20

<210> 291
<211> 21
<212> DNA
<213> Mouse

<400> 291
gatccccttt ttacagtctg c 21

<210> 292
<211> 20
<212> DNA
<213> Mouse

<400> 292
acggggttgg tactgtgtgt 20

<210> 293
<211> 20
<212> DNA
<213> Mouse

<400> 293
cacccattgt tagtgctgga 20

<210> 294
<211> 20
<212> DNA
<213> Mouse

<400> 294
acggggttgg tactgtgtgt 20

<210> 295
<211> 20
<212> DNA
<213> Mouse

<400> 295
cacacaccca cccattgtta 20

<210> 296
<211> 20
<212> DNA
<213> Mouse

<400> 296
tgcattggcc agactagaaa 20

<210> 297
<211> 19
<212> DNA
<213> Mouse

<400> 297
cggctgggct atgacctat 19

<210> 298
<211> 20
<212> DNA
<213> Mouse

<400> 298
tgcattggcc agactagaaa 20

<210> 299
<211> 20

<212> DNA
<213> Mouse

<400> 299
cggctgggct atgacctatt 20

<210> 300
<211> 20
<212> DNA
<213> Mouse

<400> 300
gttctgcagc atgatgtcgt 20

<210> 301
<211> 20
<212> DNA
<213> Mouse

<400> 301
ggcagttgtg actctgttgc 20

<210> 302
<211> 20
<212> DNA
<213> Mouse

<400> 302
gttctgcagc atgatgtcgt 20

<210> 303
<211> 20
<212> DNA
<213> Mouse

<400> 303
ctgcaggcag ttgtgactct 20

<210> 304
<211> 20
<212> DNA
<213> Mouse

<400> 304
ccatcctttt tgcctgtctt 20

<210> 305
<211> 20
<212> DNA
<213> Mouse

<400> 305
tctggaggaa catgtgatgg 20

<210> 306
<211> 20
<212> DNA
<213> Mouse

<400> 306
caccatcctt tttgcctgtc 20

<210> 307
<211> 19
<212> DNA
<213> Mouse

<400> 307
gaacatgtga tggggcaac 19

<210> 308
<211> 19
<212> DNA
<213> Mouse

<400> 308
caaagcagca ggaggagtg 19

<210> 309
<211> 20
<212> DNA
<213> Mouse

<400> 309
aaatgtactg gccaggcaac 20

<210> 310
<211> 20
<212> DNA
<213> Mouse

<400> 310
agtgctagac ccagcaccag 20

<210> 311
<211> 20
<212> DNA
<213> Mouse

<400> 311
aaatgtactg gccaggcaac 20

<210> 312
<211> 20
<212> DNA
<213> Mouse

<400> 312
gcactgacca gtctgtcacc 20

<210> 313
<211> 20
<212> DNA
<213> Mouse

<400> 313
gtccccagag aaaagcacag 20

<210> 314
<211> 20
<212> DNA
<213> Mouse

<400> 314
cagtctgtca ccacctctgg 20

<210> 315
<211> 20

<212> DNA
<213> Mouse

<400> 315
cagtggtccc cagagaaaag 20

<210> 316
<211> 20
<212> DNA
<213> Mouse

<400> 316
tactattcgg ggcttgttgg 20

<210> 317
<211> 20
<212> DNA
<213> Mouse

<400> 317
gcagcactat gtgcctggta 20

<210> 318
<211> 20
<212> DNA
<213> Mouse

<400> 318
tactattcgg ggcttgttgg 20

<210> 319
<211> 20
<212> DNA
<213> Mouse

<400> 319
gcctggtatt tgatcgcttt 20

<210> 320
<211> 20
<212> DNA
<213> Mouse

<400> 320
gctcagctag ggatggagaa 20

<210> 321
<211> 20
<212> DNA
<213> Mouse

<400> 321
cagctcaggg acacaatgaa 20

<210> 322
<211> 20
<212> DNA
<213> Mouse

<400> 322
tcctacaggc tagggctcag 20

<210> 323
<211> 20
<212> DNA
<213> Mouse

<400> 323
cagctcaggg acacaatgaa 20

<210> 324
<211> 20
<212> DNA
<213> Mouse

<400> 324
gggactgatg tgtggcttgt 20

<210> 325
<211> 20
<212> DNA
<213> Mouse

<400> 325
aggcgtccca ggaatagaag 20

<210> 326
<211> 21
<212> DNA
<213> Mouse

<400> 326
ggactgatgt gtggcttgtt t 21

<210> 327
<211> 20
<212> DNA
<213> Mouse

<400> 327
aggcgtccca ggaatagaag 20

<210> 328
<211> 20
<212> DNA
<213> Mouse

<400> 328
tgtttctgtt ctggtggctg 20

<210> 329
<211> 20
<212> DNA
<213> Mouse

<400> 329
atctgcaggc aggatcagac 20

<210> 330
<211> 20
<212> DNA
<213> Mouse

<400> 330
ctcagtggtg ggtgacagtg 20

<210> 331
<211> 20

<212> DNA
<213> Mouse

<400> 331
atctgcaggc aggatcagac 20

<210> 332
<211> 20
<212> DNA
<213> Mouse

<400> 332
acacacagta ccaaccccgt 20

<210> 333
<211> 20
<212> DNA
<213> Mouse

<400> 333
cctgtggtga tcaagaagca 20

<210> 334
<211> 20
<212> DNA
<213> Mouse

<400> 334
tgcttcttga tcaccacagg 20

<210> 335
<211> 20
<212> DNA
<213> Mouse

<400> 335
gcaacagagt cacaactgcc 20

<210> 336
<211> 20
<212> DNA
<213> Mouse

<400> 336
acacacagta ccaaccccgt 20

<210> 337
<211> 20
<212> DNA
<213> Mouse

<400> 337
gcaacagagt cacaactgcc 20

<210> 338
<211> 20
<212> DNA
<213> Mouse

<400> 338
gggtttatgt ggcaagcact 20

<210> 339
<211> 20
<212> DNA
<213> Mouse

<400> 339
actccatttg ccttttgtgg 20

<210> 340
<211> 20
<212> DNA
<213> Mouse

<400> 340
cgctacttcg cttttatccg 20

<210> 341
<211> 20
<212> DNA
<213> Mouse

<400> 341
atgatgacgt acgacgacga 20

<210> 342
<211> 21
<212> DNA
<213> Mouse

<400> 342
gaaaacaatc ggggagaagt c 21

<210> 343
<211> 20
<212> DNA
<213> Mouse

<400> 343
tgaaattatc acacgccagg 20

<210> 344
<211> 20
<212> DNA
<213> Mouse

<400> 344
agtgagaggc ccagtctcaa 20

<210> 345
<211> 20
<212> DNA
<213> Mouse

<400> 345
gatctgatgc cctcttctgc 20

<210> 346
<211> 20
<212> DNA
<213> Mouse

<400> 346
gctagccttg aagccaacac 20

<210> 347
<211> 20

<212> DNA
<213> Mouse

<400> 347
tgaacagcat gcttacccag 20

<210> 348
<211> 20
<212> DNA
<213> Mouse

<400> 348
tccctagagg cctgtctgtc 20

<210> 349
<211> 20
<212> DNA
<213> Mouse

<400> 349
tcgtctcgga gcctcttcta 20

<210> 350
<211> 20
<212> DNA
<213> Mouse

<400> 350
gatagtccct tagccagccc 20

<210> 351
<211> 20
<212> DNA
<213> Mouse

<400> 351
gccatagctc ctcactgctc 20

<210> 352
<211> 20
<212> DNA
<213> Mouse

<400> 352
cagagtgggc tctggtcttc 20

<210> 353
<211> 20
<212> DNA
<213> Mouse

<400> 353
ttgtgttcag atgctcctgc 20

<210> 354
<211> 20
<212> DNA
<213> Mouse

<400> 354
ttatttctgt gctagccgcc 20

<210> 355
<211> 20
<212> DNA
<213> Mouse

<400> 355
atcaagtcaa cgtccccaag 20

<210> 356
<211> 20
<212> DNA
<213> Mouse

<400> 356
acctggcctg tgctaatctc 20

<210> 357
<211> 20
<212> DNA
<213> Mouse

<400> 357
gcaccaaccc taagaaagca 20

<210> 358
<211> 22
<212> DNA
<213> Mouse

<400> 358
tcaggctaac ctcaaactca ca 22

<210> 359
<211> 27
<212> DNA
<213> Mouse

<400> 359
aaagaaaaga aaagaaaaag tcagaca 27

<210> 360
<211> 20
<212> DNA
<213> Mouse

<400> 360
cccagaactc catcctcaaa 20

<210> 361
<211> 20
<212> DNA
<213> Mouse

<400> 361
cccaacctgt ggtcagctat 20

<210> 362
<211> 20
<212> DNA
<213> Mouse

<400> 362
ggggcaggtg ggtaataagt 20

<210> 363
<211> 20

<212> DNA
<213> Mouse

<400> 363
caaaagccca actccttgag 20

<210> 364
<211> 20
<212> DNA
<213> Mouse

<400> 364
gctcagtggg taagagcacc 20

<210> 365
<211> 20
<212> DNA
<213> Mouse

<400> 365
ctaccctgcc gctaatctca 20

<210> 366
<211> 20
<212> DNA
<213> Mouse

<400> 366
cagttagcac cccaccctaa 20

<210> 367
<211> 20
<212> DNA
<213> Mouse

<400> 367
tctgcacctc tgttcacctg 20

<210> 368
<211> 20
<212> DNA
<213> Mouse

<400> 368
acctctaggg tttacgggga 20

<210> 369
<211> 20
<212> DNA
<213> Mouse

<400> 369
cctcaggtag tgcaagctcc 20

<210> 370
<211> 20
<212> DNA
<213> Mouse

<400> 370
tcagttacca agggtttcgg 20

<210> 371
<211> 20
<212> DNA
<213> Mouse

<400> 371
ataggttgtc acaggccagg 20

<210> 372
<211> 20
<212> DNA
<213> Mouse

<400> 372
tcagttacca agggtttcgg 20

<210> 373
<211> 20
<212> DNA
<213> Mouse

<400> 373
ataggttgtc acaggccagg 20

<210> 374
<211> 20
<212> DNA
<213> Mouse

<400> 374
gtggttgctg ggatttgaac 20

<210> 375
<211> 20
<212> DNA
<213> Mouse

<400> 375
caagcaacca aacaaccaaa 20

<210> 376
<211> 20
<212> DNA
<213> Mouse

<400> 376
tccggaggac cataaatctg 20

<210> 377
<211> 20
<212> DNA
<213> Mouse

<400> 377
cacagtccca gtcattccct 20

<210> 378
<211> 20
<212> DNA
<213> Mouse

<400> 378
gtcccaaaag ctagcacagg 20

<210> 379
<211> 20

<212> DNA
<213> Mouse

<400> 379
tcatgagcca ccatgtgatt 20

<210> 380
<211> 20
<212> DNA
<213> Mouse

<400> 380
gaccttcgga agagcagttg 20

<210> 381
<211> 20
<212> DNA
<213> Mouse

<400> 381
agtgtgtgtc gccatatcca 20

<210> 382
<211> 20
<212> DNA
<213> Mouse

<400> 382
cctactctct ctccccgctt 20

<210> 383
<211> 20
<212> DNA
<213> Mouse

<400> 383
ggaaaatgtt tggccttgaa 20

<210> 384
<211> 20
<212> DNA
<213> Mouse

<400> 384
ctggagtgaa aggcaggaag 20

<210> 385
<211> 20
<212> DNA
<213> Mouse

<400> 385
aggcggcacc atatgaataa 20

<210> 386
<211> 21
<212> DNA
<213> Mouse

<400> 386
tgagagtggg aattctgttc a 21

<210> 387
<211> 20
<212> DNA
<213> Mouse

<400> 387
ggatgtaatt ggtggcaagg 20

<210> 388
<211> 20
<212> DNA
<213> Mouse

<400> 388
ctgttggagg aggtggccta 20

<210> 389
<211> 21
<212> DNA
<213> Mouse

<400> 389
tgcttgtatg tttttcctcg t 21

<210> 390
<211> 20
<212> DNA
<213> Mouse

<400> 390
tgagagtgcc ctcctctttg 20

<210> 391
<211> 18
<212> DNA
<213> Mouse

<400> 391
gaacccctga ccccagac 18

<210> 392
<211> 22
<212> DNA
<213> Mouse

<400> 392
tgaagtgcag atttttacat gg 22

<210> 393
<211> 20
<212> DNA
<213> Mouse

<400> 393
gttttggggt ggaaaaggat 20

<210> 394
<211> 20
<212> DNA
<213> Mouse

<400> 394
ccgtcgacat ttaggtgaca 20

<210> 395
<211> 20

<212> DNA
<213> Mouse

<400> 395
gatactgggg tggtgggtaa 20

<210> 396
<211> 20
<212> DNA
<213> Mouse

<400> 396
ccgtcgacat ttaggtgaca 20

<210> 397
<211> 20
<212> DNA
<213> Mouse

<400> 397
cgtcccagct gtgtaactga 20

<210> 398
<211> 21
<212> DNA
<213> Mouse

<400> 398
ggaagcaaat gctccactaa a 21

<210> 399
<211> 20
<212> DNA
<213> Mouse

<400> 399
tatccctagc cccttgtgtg 20

<210> 400
<211> 20
<212> DNA
<213> Mouse

<400> 400
ccgtcgacat ttaggtgaca 20

<210> 401
<211> 20
<212> DNA
<213> Mouse

<400> 401
gggtcctgtt ggtagtgacc 20

<210> 402
<211> 20
<212> DNA
<213> Mouse

<400> 402
tataagcagc ccctcattgg 20

<210> 403
<211> 20
<212> DNA
<213> Mouse

<400> 403
caggccagac actgcttaca 20

<210> 404
<211> 20
<212> DNA
<213> Mouse

<400> 404
ccttgggatc tggtgtgact 20

<210> 405
<211> 20
<212> DNA
<213> Mouse

<400> 405
tgggtttaga gtacggctgg 20

<210> 406
<211> 20
<212> DNA
<213> Mouse

<400> 406
acccatttcc taatcccctg 20

<210> 407
<211> 20
<212> DNA
<213> Mouse

<400> 407
atctctccag cccctctcag 20

<210> 408
<211> 20
<212> DNA
<213> Mouse

<400> 408
gggctgggaa ttgaacctat 20

<210> 409
<211> 20
<212> DNA
<213> Mouse

<400> 409
tgaatccctt acagccttgc 20

<210> 410
<211> 20
<212> DNA
<213> Mouse

<400> 410
gccccataaa atccactcct 20

<210> 411
<211> 20

<212> DNA
<213> Mouse

<400> 411
gctccggaag gctagaagat 20

<210> 412
<211> 20
<212> DNA
<213> Mouse

<400> 412
ggtttgggag tgttaggcaa 20

<210> 413
<211> 20
<212> DNA
<213> Mouse

<400> 413
actcagttgg cctctcctca 20

<210> 414
<211> 19
<212> DNA
<213> Mouse

<400> 414
acagaaatcc ctcatgcga 19

<210> 415
<211> 21
<212> DNA
<213> Mouse

<400> 415
tcagtgtgga ccagaaagtc c 21

<210> 416
<211> 22
<212> DNA
<213> Mouse

<400> 416
tctgcaagtc agctcttgat aa 22

<210> 417
<211> 23
<212> DNA
<213> Mouse

<400> 417
actcataagg gtcaagctgt ctg 23

<210> 418
<211> 20
<212> DNA
<213> Mouse

<400> 418
tctccccttt taccactccc 20

<210> 419
<211> 20
<212> DNA
<213> Mouse

<400> 419
gcaaggagtc aaaaacagca 20

<210> 420
<211> 20
<212> DNA
<213> Mouse

<400> 420
gctagttggg gaacaaacca 20

<210> 421
<211> 20
<212> DNA
<213> Mouse

<400> 421
actgcaaatg tccaactcca 20

<210> 422
<211> 20
<212> DNA
<213> Mouse

<400> 422
cagttacaca gctgggacga 20

<210> 423
<211> 20
<212> DNA
<213> Mouse

<400> 423
gcaagagcct agcaatccac 20

<210> 424
<211> 20
<212> DNA
<213> Mouse

<400> 424
cagtttagca ccccacccta 20

<210> 425
<211> 20
<212> DNA
<213> Mouse

<400> 425
tctgcacctc tgttcacctg 20

<210> 426
<211> 20
<212> DNA
<213> Mouse

<400> 426
gggttccact tgatgctgat 20

<210> 427
<211> 20

<212> DNA
<213> Mouse

<400> 427
tggtctgttt cctggagctt 20

<210> 428
<211> 21
<212> DNA
<213> Mouse

<400> 428
tgtagggaat gtttctgcac c 21

<210> 429
<211> 20
<212> DNA
<213> Mouse

<400> 429
acatggaaca ggattctggc 20

<210> 430
<211> 20
<212> DNA
<213> Mouse

<400> 430
gcaggcaaac agacagacaa 20

<210> 431
<211> 20
<212> DNA
<213> Mouse

<400> 431
atgggggatc ccttactgac 20

<210> 432
<211> 20
<212> DNA
<213> Mouse

<400> 432
cggtcaggag tagtgtgggt 20

<210> 433
<211> 20
<212> DNA
<213> Mouse

<400> 433
cagcagctga tattgaggca 20

<210> 434
<211> 22
<212> DNA
<213> Mouse

<400> 434
aatgatgaag tgtcagcctc ag 22

<210> 435
<211> 20
<212> DNA
<213> Mouse

<400> 435
caacagaact caaagcctgg 20

<210> 436
<211> 20
<212> DNA
<213> Mouse

<400> 436
agcaggcaca ggtctcttgt 20

<210> 437
<211> 20
<212> DNA
<213> Mouse

<400> 437
aagaacagga cagtggtggg 20

<210> 438
<211> 20
<212> DNA
<213> Mouse

<400> 438
cagcgattgg ctcttctctt 20

<210> 439
<211> 20
<212> DNA
<213> Mouse

<400> 439
ggggcttcct ttctgaggta 20

<210> 440
<211> 20
<212> DNA
<213> Mouse

<400> 440
agctcaggtc cagcttggta 20

<210> 441
<211> 20
<212> DNA
<213> Mouse

<400> 441
attttcccct cctgcttctc 20

<210> 442
<211> 20
<212> DNA
<213> Mouse

<400> 442
ccaagcctct gctggttatc 20

<210> 443
<211> 20

<212> DNA
<213> Mouse

<400> 443
tgagggtgga gaatggaaag 20

<210> 444
<211> 20
<212> DNA
<213> Mouse

<400> 444
gccccataaa atccactcct 20

<210> 445
<211> 20
<212> DNA
<213> Mouse

<400> 445
ttgcctaaca ctcccaaacc 20

<210> 446
<211> 20
<212> DNA
<213> Mouse

<400> 446
cagttacaca gctgggacga 20

<210> 447
<211> 20
<212> DNA
<213> Mouse

<400> 447
gcaagagcct agcaatccac 20

<210> 448
<211> 20
<212> DNA
<213> Mouse

<400> 448
cagcaccttc ctctggtctc 20

<210> 449
<211> 20
<212> DNA
<213> Mouse

<400> 449
tgtctccaga ggttctgcct 20

<210> 450
<211> 24
<212> DNA
<213> Mouse

<400> 450
tggtggtgta atactattcc tttg 24

<210> 451
<211> 26
<212> DNA
<213> Mouse

<400> 451
tctttaattt ttggcttttt gataca 26

<210> 452
<211> 20
<212> DNA
<213> Mouse

<400> 452
cagctgtgtg catgttgacc 20

<210> 453
<211> 20
<212> DNA
<213> Mouse

<400> 453
catcatgaag actcagggca 20

<210> 454
<211> 20
<212> DNA
<213> Mouse

<400> 454
gtccacacct ggcttttgtt 20

<210> 455
<211> 20
<212> DNA
<213> Mouse

<400> 455
cagcactcag tgaggttcca 20

<210> 456
<211> 20
<212> DNA
<213> Mouse

<400> 456
atgtaatgga agggctgctg 20

<210> 457
<211> 20
<212> DNA
<213> Mouse

<400> 457
cagcactcag tgaggttcca 20

<210> 458
<211> 21
<212> DNA
<213> Mouse

<400> 458
aaacaggcat gaaactcagg a 21

<210> 459
<211> 20

<212> DNA
<213> Mouse

<400> 459
gggtatcatt gtcacctcca 20

<210> 460
<211> 20
<212> DNA
<213> Mouse

<400> 460
cacaggccaa gttgttgttg 20

<210> 461
<211> 20
<212> DNA
<213> Mouse

<400> 461
caggggacct tctgaatgat 20

<210> 462
<211> 20
<212> DNA
<213> Mouse

<400> 462
agctcaggtc cagcttggta 20

<210> 463
<211> 20
<212> DNA
<213> Mouse

<400> 463
accacaaaat tttcccctcc 20

<210> 464
<211> 20
<212> DNA
<213> Mouse

<400> 464
cgggacctaa aactggacaa 20

<210> 465
<211> 20
<212> DNA
<213> Mouse

<400> 465
tggggacagt taccaggaag 20

<210> 466
<211> 20
<212> DNA
<213> Mouse

<400> 466
ccggaggacc ataaatctga 20

<210> 467
<211> 20
<212> DNA
<213> Mouse

<400> 467
cctcaaaaac aagcctgagc 20

<210> 468
<211> 22
<212> DNA
<213> Mouse

<400> 468
ccttcagaaa tgtgtttgga ca 22

<210> 469
<211> 20
<212> DNA
<213> Mouse

<400> 469
tcctgagttc aaatcccagc 20

<210> 470
<211> 20
<212> DNA
<213> Mouse

<400> 470
ctttccattc tccaccctca 20

<210> 471
<211> 20
<212> DNA
<213> Mouse

<400> 471
aggtcctagg gagaggtcca 20

<210> 472
<211> 20
<212> DNA
<213> Mouse

<400> 472
aggcctaccc aaggacatct 20

<210> 473
<211> 20
<212> DNA
<213> Mouse

<400> 473
gcagtgagct gcagagtttg 20

<210> 474
<211> 20
<212> DNA
<213> Mouse

<400> 474
agacacccta ggtcctgctg 20

<210> 475
<211> 22

<212> DNA
<213> Mouse

<400> 475
tgatctttcc aaacgcataa ga 22

<210> 476
<211> 20
<212> DNA
<213> Mouse

<400> 476
gcaagcaacc tgaacatgaa 20

<210> 477
<211> 20
<212> DNA
<213> Mouse

<400> 477
gcttacgatg gtcgtgaggt 20

<210> 478
<211> 20
<212> DNA
<213> Mouse

<400> 478
acatgcctgc ctatctttgc 20

<210> 479
<211> 20
<212> DNA
<213> Mouse

<400> 479
ggaacctgtt ttccatggtg 20

<210> 480
<211> 20
<212> DNA
<213> Mouse

<400> 480
accttgttcc tggtgtgagc 20

<210> 481
<211> 20
<212> DNA
<213> Mouse

<400> 481
tagctgggac gtggtatggt 20

<210> 482
<211> 20
<212> DNA
<213> Mouse

<400> 482
ccatgggaga ccagaaggta 20

<210> 483
<211> 20
<212> DNA
<213> Mouse

<400> 483
tgagtgtcct ctgcctgatg 20

<210> 484
<211> 20
<212> DNA
<213> Mouse

<400> 484
gcgctgacat cctcctatgt 20

<210> 485
<211> 20
<212> DNA
<213> Mouse

<400> 485
cccactatgg tcccagagaa 20

<210> 486
<211> 20
<212> DNA
<213> Mouse

<400> 486
ttgcacgtct ttgtttcgag 20

<210> 487
<211> 24
<212> DNA
<213> Mouse

<400> 487
aaaggggaat agacctgagt agaa 24

<210> 488
<211> 20
<212> DNA
<213> Mouse

<400> 488
ccaagagtca gccttggagt 20

<210> 489
<211> 20
<212> DNA
<213> Mouse

<400> 489
ggacaggtag ctcacccaac 20

<210> 490
<211> 19
<212> DNA
<213> Mouse

<400> 490
tgccagcttt ggctatcat 19

<210> 491
<211> 20

<212> DNA
<213> Mouse

<400> 491
ttcattgtgt ccctgagctg 20

<210> 492
<211> 24
<212> DNA
<213> Mouse

<400> 492
agctttggct atcatgggtc tcag 24

<210> 493
<211> 22
<212> DNA
<213> Mouse

<400> 493
accaccgcca ctgttctcat ct 22

<210> 494
<211> 20
<212> DNA
<213> Mouse

<400> 494
tgtgggggaa gaacatagaa 20

<210> 495
<211> 22
<212> DNA
<213> Mouse

<400> 495
tgatgtgtgg cttgtttctc tt 22

<210> 496
<211> 20
<212> DNA
<213> Mouse

<400> 496
ataggtgggg agggagctaa 20

<210> 497
<211> 22
<212> DNA
<213> Mouse

<400> 497
tgatgtgtgg cttgtttctc tt 22

<210> 498
<211> 20
<212> DNA
<213> Homo sapiens

<400> 498
tgtgcctgtc acagcaactt 20

<210> 499
<211> 20
<212> DNA
<213> Homo sapiens

<400> 499
catgctagca ccgtagctga 20

<210> 500
<211> 20
<212> DNA
<213> Homo sapiens

<400> 500
ggagaccttc ccctccttct 20

<210> 501
<211> 20
<212> DNA
<213> Homo sapiens

<400> 501
gctgtagttg aagagggcgt 20

<210> 502
<211> 18
<212> DNA
<213> Homo sapiens

<400> 502
gtgcttggct tcctccag 18

<210> 503
<211> 20
<212> DNA
<213> Homo sapiens

<400> 503
caggtcgtac tccatgtcca 20

<210> 504
<211> 20
<212> DNA
<213> Homo sapiens

<400> 504
tggagtacga cctgaagctg 20

<210> 505
<211> 20
<212> DNA
<213> Homo sapiens

<400> 505
actcatcctg gccacaaaag 20

<210> 506
<211> 19
<212> DNA
<213> Homo sapiens

<400> 506
gaacaggagg acgctgagg 19

<210> 507
<211> 20

<212> DNA
<213> Homo sapiens

<400> 507
cttttgtggc caggatgagt 20

<210> 508
<211> 20
<212> DNA
<213> Homo sapiens

<400> 508
tcacctcacc tggttgtcag 20

<210> 509
<211> 20
<212> DNA
<213> Homo sapiens

<400> 509
gtacgacctg aagctgtggg 20

<210> 510
<211> 27
<212> DNA
<213> Homo sapiens

<400> 510
ggctgagatc acagggttgg gtcactc 27

<210> 511
<211> 27
<212> DNA
<213> Homo sapiens

<400> 511
ccgtgcctgt tggaagttgc ctctgcc 27

<210> 512
<211> 20
<212> DNA
<213> Mouse

<400> 512
aattcccagc aaccactcac 20

<210> 513
<211> 20
<212> DNA
<213> Mouse

<400> 513
cagacactcc agaagagggc 20

<210> 514
<211> 20
<212> DNA
<213> Mouse

<400> 514
tgactgctct tccgaaggtt 20

<210> 515
<211> 20
<212> DNA
<213> Mouse

<400> 515
tttgtggaat agccaaagcc 20

<210> 516
<211> 20
<212> DNA
<213> Mouse

<400> 516
tctctcctct cttctccccc 20

<210> 517
<211> 20
<212> DNA
<213> Mouse

<400> 517
agcagggtgc atcaccttat 20

<210> 518
<211> 20
<212> DNA
<213> Mouse

<400> 518
taggagtgcc ccataggttg 20

<210> 519
<211> 20
<212> DNA
<213> Mouse

<400> 519
tcattgtacc cagccagtca 20

<210> 520
<211> 20
<212> DNA
<213> Mouse

<400> 520
aggactgagc ctggatgaga 20

<210> 521
<211> 20
<212> DNA
<213> Mouse

<400> 521
ctgggcgttt tgttttgttt 20

<210> 522
<211> 20
<212> DNA
<213> Mouse

<400> 522
cttcctcctg cagctaccac 20

<210> 523
<211> 20

<212> DNA
<213> Mouse

<400> 523
accctgctac aacgcagact 20

<210> 524
<211> 20
<212> DNA
<213> Mouse

<400> 524
tccaaccttg acacccattt 20

<210> 525
<211> 20
<212> DNA
<213> Mouse

<400> 525
agccagggct acacagagaa 20

<210> 526
<211> 20
<212> DNA
<213> Mouse

<400> 526
ctgcttttcc tcagcaactg 20

<210> 527
<211> 20
<212> DNA
<213> Mouse

<400> 527
attcgccgtt agaagctagg 20

<210> 528
<211> 20
<212> DNA
<213> Mouse

<400> 528
aactgtacgt ggctgctggt 20

<210> 529
<211> 20
<212> DNA
<213> Mouse

<400> 529
attcgccgtt agaagctagg 20

<210> 530
<211> 20
<212> DNA
<213> Mouse

<400> 530
gccaggtgac ccttatgaaa 20

<210> 531
<211> 20
<212> DNA
<213> Mouse

<400> 531
gagagatggc agacagaggc 20

<210> 532
<211> 20
<212> DNA
<213> Mouse

<400> 532
agctctctgt ccctggtgaa 20

<210> 533
<211> 20
<212> DNA
<213> Mouse

<400> 533
tgccaaccac tagcctctct 20

<210> 534
<211> 20
<212> DNA
<213> Mouse

<400> 534
ctgaaccctc cactctcctg 20

<210> 535
<211> 20
<212> DNA
<213> Mouse

<400> 535
agccagggct acacagagaa 20

<210> 536
<211> 20
<212> DNA
<213> Mouse

<400> 536
agccagggct acacagagaa 20

<210> 537
<211> 20
<212> DNA
<213> Mouse

<400> 537
accctgctac aacgcagact 20

<210> 538
<211> 20
<212> DNA
<213> Mouse

<400> 538
gcaagtttca ggagctaggg 20

<210> 539
<211> 20

<212> DNA
<213> Mouse

<400> 539
ccccagaacc agagaccata 20

<210> 540
<211> 20
<212> DNA
<213> Mouse

<400> 540
ccccagaacc agagaccata 20

<210> 541
<211> 20
<212> DNA
<213> Mouse

<400> 541
ctaggggact ctgccaagtg 20

<210> 542
<211> 20
<212> DNA
<213> Mouse

<400> 542
caagacaccc agtcccaact 20

<210> 543
<211> 20
<212> DNA
<213> Mouse

<400> 543
tacttcccct ttcccgaact 20

<210> 544
<211> 20
<212> DNA
<213> Mouse

<400> 544
tccttggtgc ttaccctcac 20

<210> 545
<211> 20
<212> DNA
<213> Mouse

<400> 545
tgttcctgag ttcacaacgc 20

<210> 546
<211> 20
<212> DNA
<213> Mouse

<400> 546
attcccagca actacatggc 20

<210> 547
<211> 20
<212> DNA
<213> Mouse

<400> 547
acatgtccac tgtggcaaaa 20

<210> 548
<211> 20
<212> DNA
<213> Mouse

<400> 548
tgtcatgagt ttgaggccag 20

<210> 549
<211> 20
<212> DNA
<213> Mouse

<400> 549
atcagacagc ccacaacctc 20

<210> 550
<211> 20
<212> DNA
<213> Mouse

<400> 550
tatgtgccac cacacctgtc 20

<210> 551
<211> 20
<212> DNA
<213> Mouse

<400> 551
gctcaaggaa ggacacacct 20

<210> 552
<211> 22
<212> DNA
<213> Mouse

<400> 552
tgctcttaac attttgagcc at 22

<210> 553
<211> 20
<212> DNA
<213> Mouse

<400> 553
gctcagcccc tgaatcaata 20

<210> 554
<211> 20
<212> DNA
<213> Mouse

<400> 554
gggatctgcc tgtcttacca 20

<210> 555
<211> 20

<212> DNA
<213> Mouse

<400> 555
ggaaggtagg gcctggtaat 20

<210> 556
<211> 20
<212> DNA
<213> Mouse

<400> 556
gctccaagat ctgtgcgatt 20

<210> 557
<211> 20
<212> DNA
<213> Mouse

<400> 557
ttagcgttag ggtgagggtg 20

<210> 558
<211> 20
<212> DNA
<213> Mouse

<400> 558
ggagactacg gacttgtggc 20

<210> 559
<211> 20
<212> DNA
<213> Mouse

<400> 559
cagttcttcc cgaaaaccac 20

<210> 560
<211> 20
<212> DNA
<213> Mouse

<400> 560
tttctgggaa ctgagatggc 20

<210> 561
<211> 20
<212> DNA
<213> Mouse

<400> 561
gttggggctg ctcatagaaa 20

<210> 562
<211> 20
<212> DNA
<213> Mouse

<400> 562
gctgtggctc tcttggagtt 20

<210> 563
<211> 20
<212> DNA
<213> Mouse

<400> 563
ctctgatttc ccacatgcct 20

<210> 564
<211> 20
<212> DNA
<213> Mouse

<400> 564
aagagggagc actgaggaca 20

<210> 565
<211> 20
<212> DNA
<213> Mouse

<400> 565
cagcagcaaa tgacctttca 20

<210> 566
<211> 20
<212> DNA
<213> Mouse

<400> 566
gaggcaggca gatttctgag 20

<210> 567
<211> 20
<212> DNA
<213> Mouse

<400> 567
gtttcacatg ttgtggtggc 20

<210> 568
<211> 20
<212> DNA
<213> Mouse

<400> 568
gggacctttg ggatagcatt 20

<210> 569
<211> 20
<212> DNA
<213> Mouse

<400> 569
tcagacatct ctggcctcct 20

<210> 570
<211> 20
<212> DNA
<213> Mouse

<400> 570
ttcactaagt tgcccaggct 20

<210> 571
<211> 22

<212> DNA
<213> Mouse

<400> 571
tgccttttc tcacattgtc tc 22

<210> 572
<211> 20
<212> DNA
<213> Mouse

<400> 572
ttagaagcag aggcagaggc 20

<210> 573
<211> 20
<212> DNA
<213> Mouse

<400> 573
gacctttgga agagcagtcg 20

<210> 574
<211> 20
<212> DNA
<213> Mouse

<400> 574
tggcagctca caatgtcttt 20

<210> 575
<211> 20
<212> DNA
<213> Mouse

<400> 575
ggtgtggtgt aggggaagaa 20

<210> 576
<211> 22
<212> DNA
<213> Mouse

<400> 576
tttcaactgc aaacacaaac ag 22

<210> 577
<211> 19
<212> DNA
<213> Mouse

<400> 577
agggccaagg aaggagaat 19

<210> 578
<211> 24
<212> DNA
<213> Mouse

<400> 578
gcaaatatat agggtaccga gctg 24

<210> 579
<211> 20
<212> DNA
<213> Mouse

<400> 579
cagattctcc agctgtcagg 20

<210> 580
<211> 19
<212> DNA
<213> Mouse

<400> 580
ctgtgtttcc gcaccaagt 19

<210> 581
<211> 20
<212> DNA
<213> Mouse

<400> 581
ctgcccgtcc ttatcttctg 20

<210> 582
<211> 20
<212> DNA
<213> Mouse

<400> 582
acgcacgctc actcatacac 20

<210> 583
<211> 20
<212> DNA
<213> Mouse

<400> 583
cagcagaggt gatgggttct 20

<210> 584
<211> 22
<212> DNA
<213> Mouse

<400> 584
ttgtcacaca gtggttaaat gc 22

<210> 585
<211> 20
<212> DNA
<213> Mouse

<400> 585
tagaaccgtg gctgaggact 20

<210> 586
<211> 24
<212> DNA
<213> Mouse

<400> 586
ccgtaagata tgaaagaact tgga 24

<210> 587
<211> 20

<212> DNA
<213> Mouse

<400> 587
taatcctggc ttagcgcttg 20

<210> 588
<211> 20
<212> DNA
<213> Mouse

<400> 588
tagaaagcac aggggacagg 20

<210> 589
<211> 20
<212> DNA
<213> Mouse

<400> 589
ccttcctcgt ctgagctgtt 20

<210> 590
<211> 20
<212> DNA
<213> Mouse

<400> 590
ttgggacgtg acctgagaat 20

<210> 591
<211> 20
<212> DNA
<213> Mouse

<400> 591
tatgtgtctg gccgttgttc 20

<210> 592
<211> 19
<212> DNA
<213> Mouse

<400> 592
gatgtgggtg caggtgaag 19

<210> 593
<211> 20
<212> DNA
<213> Mouse

<400> 593
ccccttctgg agtgtctgaa 20

<210> 594
<211> 21
<212> DNA
<213> Mouse

<400> 594
tctaggcagg gctacctttt t 21

<210> 595
<211> 19
<212> DNA
<213> Mouse

<400> 595
gctgagcagc ctctagcaa 19

<210> 596
<211> 20
<212> DNA
<213> Mouse

<400> 596
accatggctt ttcccagtaa 20

<210> 597
<211> 20
<212> DNA
<213> Mouse

<400> 597
ctgtgccttt ggtgatcaga 20

<210> 598
<211> 20
<212> DNA
<213> Mouse

<400> 598
tgtggcactc tacggcataa 20

<210> 599
<211> 23
<212> DNA
<213> Mouse

<400> 599
tgcatcacta ttaagcctca acc 23

<210> 600
<211> 23
<212> DNA
<213> Mouse

<400> 600
aagaatttgc aaagactgtg aga 23

<210> 601
<211> 20
<212> DNA
<213> Mouse

<400> 601
ctggaccttt ggaagagcag 20

<210> 602
<211> 20
<212> DNA
<213> Mouse

<400> 602
ggtggctcaa accatccata 20

<210> 603
<211> 20

<212> DNA
<213> Mouse

<400> 603
gagggcaatg agcaaaatgt 20

<210> 604
<211> 20
<212> DNA
<213> Mouse

<400> 604
ggtcctgtct ctggttcagg 20

<210> 605
<211> 20
<212> DNA
<213> Mouse

<400> 605
taacacccac atcaggcaac 20

<210> 606
<211> 22
<212> DNA
<213> Mouse

<400> 606
tttcatttcc tggtgttcct tt 22

<210> 607
<211> 20
<212> DNA
<213> Mouse

<400> 607
aaacacaggc ggaacgatag 20

<210> 608
<211> 20
<212> DNA
<213> Mouse

<400> 608
ctatcgttcc gcctgtgttt 20

<210> 609
<211> 21
<212> DNA
<213> Mouse

<400> 609
aaggaagagg atggagaaag a 21

<210> 610
<211> 20
<212> DNA
<213> Mouse

<400> 610
cgggtcttaa tggagcagag 20

<210> 611
<211> 20
<212> DNA
<213> Mouse

<400> 611
tcctccccag ttacctagca 20

<210> 612
<211> 19
<212> DNA
<213> Mouse

<400> 612
cagcaggcaa gatgacctc 19

<210> 613
<211> 20
<212> DNA
<213> Mouse

<400> 613
gtccctcacc agccatgtta 20

<210> 614
<211> 20
<212> DNA
<213> Mouse

<400> 614
agcctgggct aagttgtgtg 20

<210> 615
<211> 20
<212> DNA
<213> Mouse

<400> 615
tatgggccaa tgttgttcct 20

<210> 616
<211> 20
<212> DNA
<213> Mouse

<400> 616
atggtggctc acaaccatct 20

<210> 617
<211> 20
<212> DNA
<213> Mouse

<400> 617
ttgtcctctg attgcagcat 20

<210> 618
<211> 20
<212> DNA
<213> Mouse

<400> 618
cttgggtcat caggctttgt 20

<210> 619
<211> 20

<212> DNA
<213> Mouse

<400> 619
aagctgccct gctctctcta 20

<210> 620
<211> 20
<212> DNA
<213> Mouse

<400> 620
atgctcagcc tgctttgttt 20

<210> 621
<211> 20
<212> DNA
<213> Mouse

<400> 621
gctgatagcc ctgggttcta 20

<210> 622
<211> 21
<212> DNA
<213> Mouse

<400> 622
tgtacgcaca aattgacttg c 21

<210> 623
<211> 21
<212> DNA
<213> Mouse

<400> 623
gaatccacat tgcaaagcct a 21

<210> 624
<211> 20
<212> DNA
<213> Mouse

<400> 624
cacaggcaaa tgaagggaag 20

<210> 625
<211> 20
<212> DNA
<213> Mouse

<400> 625
ccagacttct ccagctctcc 20

<210> 626
<211> 21
<212> DNA
<213> Mouse

<400> 626
tcctcgagag gctctaggtt t 21

<210> 627
<211> 20
<212> DNA
<213> Mouse

<400> 627
tgcctagtca accacaggag 20

<210> 628
<211> 21
<212> DNA
<213> Mouse

<400> 628
cctgtggttg actaggcaga a 21

<210> 629
<211> 20
<212> DNA
<213> Mouse

<400> 629
gcctgatagc ctggaataca 20

<210> 630
<211> 20
<212> DNA
<213> Mouse

<400> 630
aaagggatgt gtggcgtaag 20

<210> 631
<211> 20
<212> DNA
<213> Mouse

<400> 631
caaaacccaa ccttctcagc 20

<210> 632
<211> 20
<212> DNA
<213> Mouse

<400> 632
tgcactgacc gtgatagagg 20

<210> 633
<211> 20
<212> DNA
<213> Mouse

<400> 633
cggtgtagct ctggctgtct 20

<210> 634
<211> 20
<212> DNA
<213> Mouse

<400> 634
catctcacca actcgcactt 20

<210> 635
<211> 21

<212> DNA
<213> Mouse

<400> 635
tttctgggaa caaagaggct a 21

<210> 636
<211> 20
<212> DNA
<213> Mouse

<400> 636
gaacccaagt gttggggtaa 20

<210> 637
<211> 20
<212> DNA
<213> Mouse

<400> 637
tggaagccca tctgtctctt 20

<210> 638
<211> 20
<212> DNA
<213> Mouse

<400> 638
aaatgcaagt gggtgcttct 20

<210> 639
<211> 19
<212> DNA
<213> Mouse

<400> 639
ccagaagagg gcgtcagat 19

<210> 640
<211> 20
<212> DNA
<213> Mouse

<400> 640
ggtgtgcacc accatattca 20

<210> 641
<211> 21
<212> DNA
<213> Mouse

<400> 641
gggaattatc agccaaaaag c 21

<210> 642
<211> 20
<212> DNA
<213> Mouse

<400> 642
gcccaactga aagctcaact 20

<210> 643
<211> 21
<212> DNA
<213> Mouse

<400> 643
ggaaggggga taacaattga a 21

<210> 644
<211> 23
<212> DNA
<213> Mouse

<400> 644
tgctaatttc aagcacagtg aga 23

<210> 645
<211> 20
<212> DNA
<213> Mouse

<400> 645
agcttgacac cttgacagca 20

<210> 646
<211> 20
<212> DNA
<213> Mouse

<400> 646
aacctgcaga gaggagacca 20

<210> 647
<211> 20
<212> DNA
<213> Mouse

<400> 647
ctccaagggg aggactcatt 20

<210> 648
<211> 24
<212> DNA
<213> Mouse

<400> 648
ttcaattgag tttctctcct ctga 24

<210> 649
<211> 20
<212> DNA
<213> Mouse

<400> 649
tgcaggacca agaagtaggc 20

<210> 650
<211> 20
<212> DNA
<213> Mouse

<400> 650
cgagatctga tgccctcttc 20

<210> 651
<211> 20

<212> DNA
<213> Mouse

<400> 651
tgctgagagc agaaaaggaa 20

<210> 652
<211> 166
<212> DNA
<213> Mouse

<400> 652
gcagtgagct gcagagtttg cagaatgagg gcactctaaa ctcatcaagt gaggaggccc 60
ttccctcaca ctccagatgg ctgataggtg gcattacatg gtccancgcg cgcacgcgct 120
cagatgcaat ctccacattc ataaccagat gtccttgggt aggcct 166